Biochemical and Structural reconstitution of the outer Kinetochore of *Drosophila melanogaster*

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

Dr. rer. nat.

der Fakultät für Biologie an der

UNIVERSITÄT DUISBURG-ESSEN

vorgelegt von

Yahui Liu

aus Wuhan, China

durchgeführt am

MAX PLANCK INSTITUT FÜR MOLEKULARE PHYSIOLOGIE

Abteilung für Mechanistische Zellbiolgie

June 2017

Die der vorliegenden Arbeit zugrunde liegenden Experimente wurden am Max Planck Institut fur Molekulare Physiologie in der Abteilung für mechanistische Zellbiologie durchgeführt.

1. Gutachter: Prof. Dr. Andrea Musacchio
2. Gutachter: Prof. Dr. Hemmo Meyer

Vorsitzender des Prüfungsausschusses: Prof. Dr. Shriley Knauer

Tag der mündlichen Prüfung: 10. 10. 2017

我要将这个论文献给我的父亲母亲，他们多年来的默默支持，一直是我前行的动力。

献给我的大姐刘亚美和姐夫张景，献给小弟刘亚青和弟媳吴真真，大家庭的温暖一直激励着我。

献给家庭的两个小宝贝：祺祺和贝贝。多年以后，你们如果能读到这段文字，会不会很激动？

最后献给覃蕾：ONE WORLD, ONE LOVER

Abstract

Dividing cells go through a series of events, named the cell cycle, to generate two new cells that are genetically equal. Faithful sister chromatid separation to the arising daughter cells is a crucial step in the cell cycle. To achieve that, kinetochores, multi-protein complexes that assemble on centromeric chromatin, act as a bridge between microtubules and chromatids. In human and yeast, kinetochores consist of two major parts: the constitutive centromere-associated network (CCAN, the inner kinetochore) and the Knl1 complex, Mis12 complex, Ndc80 complex (KMN) network. The CCAN contains 16 subunits and links the kinetochore to CENP-A. The KMN network contains 10 subunits and is responsible for the binding of microtubules and the recruitment of spindle assembly checkpoint (SAC) components. The popular model organisms *Drosophila melanogaster* and related organisms have a significantly simplified kinetochore: they lost all CCAN components with the exception of CENP-C and lost two outer kinetochore subunits (one Mis12 complex (Mis12C) subunit and one Knl1 complex subunit, Zwint). In addition, there are two homologues of the Mis12C subunit Nnf1 (Nnf1a and Nnf1b). The simplified *Drosophila* kinetochore thus forms an ideal model to study conserved structural and functional properties of kinetochores. During my doctoral work, I studied the *Drosophila* KMN network using an *in vitro* biochemical reconstitution approach. My reconstitution experiments indicate that there are two distinct three-subunit Mis12 complexes in *Drosophila*, DmMis12a and DmMis12b complexes. Despite the absent subunit, negative stain electron microscopy shows that the DmMis12C has an extended polarized structures, comparable to its human and yeast counterparts. Biochemical studies combined with cross-linking-mass spectrometry (XL-MS) proved that Mis12 and Nnf1 form the backbone of Mis12C in both DmMis12a and

DmMis12b. Using deletion experiments, I proved that both the C-terminal of Knl1 and the N-terminal of CENP-C bind to DmMis12C. Furthermore, I identified the motifs responsible for the binding between CENP-C and Mis12C in Drosophila: the conserved 3 phenylalanine residues at the N-terminal of Mis12 (F12, F13 and F15) and the 27 amino acids at the N-terminal end of CENP-C (residues 9 to 35). In cooperation with Dr. Arsen Petrovic and Dr. Jenny Keller, we solved the high-resolution structure of the tetrameric human Mis12C in complex with the N-terminal segment of CENP-C. We provided a model for how Aurora B controls Mis12C assembling to kinetochore. We demonstrated that the Dsn1 subunit is an important substrate of Aurora B in the human Mis12C. Together with a structure-based *in silico* analysis, this led us to hypothesize that the lost subunit in *Drosophila* Mis12 complex could be Nsl1, and not Dsn1 as previously suggested. In summary, my studies provided new insights on the structural and functional properties of outer kinetochores in humans and *Drosophila*.

ZUSAMMENFASSUNG

Eine genaue Chromosomensegregation während der Mitose und der Meiose ist entscheidend für die zelluläre und organisatorische Lebensfähigkeit. Kinetochore verbinden Chromosomen mit Spindel-Mikrotubuli und sind für die Chromosomensegregation essentiell. Diese großen Proteingerüste entstehen aus dem Zentromer, einer spezialisierten Region des Chromosoms, angereichert mit der Histon-H3-Variante CENP-A. In den meisten Eukaryoten besteht der Kinetochore-Kern aus dem Zentromer-proximalen konstitutiven Zentromer-assoziierten Netzwerk (CCAN), das CENP-A bindet und 16 Untereinheiten enthält, und des Zentromer-distalen Knl1/Zwint, Mis12C, Ndc80-Komplexnetzwerk (KMN). Das KMN-Komplexnetzwerk bindet Mikrotubuli uns setzt sich aus 10 Untereinheiten zusammen. In der Fruchtfliege Drosophila melanogaster erlebte die Kinetochore bemerkenswerte Vereinfachungen. Alle CCAN-Untereinheiten, mit Ausnahme von Zentromer-Protein C (CENP-C) und zwei KMN-Untereinheiten, Dsn1 und Zwint, können in diesem Organismus nicht identifiziert werden. Zusätzlich sind zwei paralogs der KMN-Untereinheit Nnf1 (Nnf1a und Nnf1b) vorhanden. Schließlich erlebte die Spc105R-Untereinheit, homolog zu humanem Knl1/CASC5, im Vergleich zu anderen Organismen erhebliche Sequenzänderungen. Wir kombinierten die biochemische Rekonstitution mit biophysikalischen und strukturellen Methoden, um zu untersuchen, wie sich diese Veränderungen auf die Organisation des Drosophila KMN Netzwerks auswirken. Wir zeigen, dass die Nnf1a- und Nnf1b-Paralogs Untereinheiten unterschiedlicher Komplexe sind, die beide direkt mit Spc105 und mit CENP-C wechselwirken, wobei letztere eine Bindungsstelle an der Mis12-Untereinheit identifizieren. Unsere Studien beleuchten die strukturelle und funktionelle Organisation eines stark divergierenden Kinetochore-Teilchens.

Contents

List of Figures

List of Table

Abbreviations

APC	Anaphase Promoting Complex
AUC	Analytical Ultra Centrifugation
BUB1	Budding Uninhibited by Benomyl 1
BUBR1	Bub1 Related
CATD	CENP-A Targeting Domain
CCAN	Constitutive Centromere Associated Network
CD	Circular Dichroism
CENP	Centromere Protein
DNA	DesoxyriboNucleic Acid
EM	Electron Microscopy
EDTA	EthylenDiaminTetraAcetic acid
ESI	ElectroSpray Ionization
IEX	Ion Exchange chromatography
IPTG	Isopropyl-D-Thiogalactopyranosid
KMN complex	Knl1 complex, Mis12 complex, Ndc80 complex
MT	MicroTubule
MAD	Mitotic Arrest Deficient
MCC	Mitotic Checkpoint Complex
MPS1	Monopolar Spindle Protein 1
MW	Molecular Weight
PBS	Phosphate Buffered Saline
PEG	PolyEthylene Glycol
PEST-rich	Proline, Glutamatic acid, Serine, Threonine-rich
PLK1	Polo-Like Kinase-1
RNA	RiboNucleic Acid
RZZ	Rod/Zwilch/ZW10
SAC	Spindle Assembly Checkpoint
SDS-PAGE	Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis
SEC	Size Exclusion Chromatography
SLS	Static Light Scattering
WB	Western Blot
XL-MS	Cross-Linking-Mass Spectrometry
Y2H	Yeast two-Hybrid

1 INTRODUCTION

1.1 The cell cycle and Mitosis

The cycle of cell division is the process whereby the cell duplicates itself and its genetic material. The cell cycle is composed of an orderly sequence of events that are highly regulated (see Figure 1). When the cycle completes, the cell will have duplicated itself to create two identical daughter cells (Alberts et al., 2014). In general, the cell cycle consists of two main stages: Interphase and Mitosis (M phase).

Before the cell can go through M phase, there is a lot of very important preparatory work to be done. This work is done in interphase. Interphase can be divided into G_1 (Gap1) phase, S (Synthesis) phase and G_2 (Gap2) phase. G_1 phase begins after M phase is complete (see Figure 1). During this phase, cells increase in size and the checkpoint control in G_1 make sure that all the requirements for DNA synthesis are available to the cell, such as relative enzymes and signaling proteins (Albert et al., 2014; Lodish et al., 2007). In some cases, cells will stop cycling and go into a waiting stage, named G_0 phase. Cells in G_0 can reenter the cell cycle, but only under certain conditions. In S phase, DNA replication begins. When this phase is completed, the genetic information carrier (DNA) will be doubled so that every chromosome will be present in two copies as sister chromatids. The stage between S phase and M phase is G_2 phase. The cells will continue to grow in size during this phase. All the proteins that will be required for the M phase will be synthesized during G_2 phase (Norbury and Nurse, 1992).

When the cells pass through G_2, they enter into M phase, shown in Figure 1. In M phase chromosomes are segregated into two identical daughter cells. M phase can be

divided into two states: mitosis (where the sister chromatids are equally segregated) and cytokinesis (where the cell divides in two daughter cells).

Mitosis consists of five phases (see Figure 1) and prophase is the first stage of mitosis. The chromatin condenses and the sister chromatids, which by the end of mitosis will be pulled to different poles of the cell, become individualized. In the next phase of mitosis, prometaphase, microtubules nucleating from the poles of the cell start to find and capture the chromatid pairs. When microtubules from opposing poles each bind a sister chromatid, at a structure on the centromere known as the kinetochore, sister chromatids become aligned along the metaphase plate. When all of the sister chromatids have become aligned, the cell is in the next stage of mitosis, metaphase. Anaphase starts when the sister chromatids, now attached to the mitotic spindle, are pulled to opposing poles (Anaphase A). The spindle continues to elongate as the chromosomes approach the poles (Anaphase B) (Yanagida, 2014). The nuclear envelope re-forms in telophase and the cell starts to divide into two independent daughter cells (Yanagida, 2014).

Cytokinesis is the final step of cell division (see Figure 1). The typical characteristic of cytokinesis is the reorganization of the mitotic spindle that signals the assembly of an actomyosin ring between the spindle poles (Gould, 2016; Srivastava et al., 2016). Contraction of this ring promotes the formation of a cleavage furrow, which is a thin intercellular bridge that separates the daughter cells after completion of furrowing (Gould, 2016). And finally the cells divide into two.

The cell cycle is controlled by heterodimeric protein kinases, which contains a regulatory subunit (cyclins) and a catalytic subunit (cyclin-dependent kinases, or Cdks). The concentration and activity of these complexes oscillate during the cell cycle, leading to phosphorylation of diverse sets of substrates, ensuring directionality to the process, for instance by promoting DNA replication prior to sister chromatid separation (Alberts et al., 2014; Lodish et al., 2007).

The concentration of Cdk proteins is almost constant during the cell cycle, whereas cyclin proteins levels show cyclical changes that promote activation of cyclin-Cdk complexes at specific time points. There are four kinds of cyclins: 1. G_1/S-cyclins, whose concentration raises in late G_1 and decreases in S phase, activate Cdks in the late G_1, resulting in cell cycle entry; 2. S-cyclins form S-Cdk complex and their concentration increases at the beginning S phase, leading to DNA replication. Since they also controls some early mitotic events, their levels only decrease in mitosis; 3. M-cyclins, whose levels increase during the G_2 phase. They are functional at the end of G_2 and promote the G2/M transition. Since cyclins are substrate of APC/C, an ubiquitin ligase that target them for degradation by the 26S proteasome (discussed in more details in section 1.3), their levels decrease at the end of mitosis; 4. G_1-cyclin is not functionally independent and it controls the activities of the G_1-S cyclins at late G_1 (Holt et al., 2009; Lodish et al., 2007; Pavletich, 1999; Wittenberg and Reed, 2005).

There is an additional checkpoint activity in mitosis: the spindle assembly checkpoint (SAC), which is a feedback system that controls the transition from metaphase to anaphase (Walczak and Heald, 2008). It will be discussed further in section 1.3.

1.2 Microtubules and Mitotic spindle

Microtubules are polarized protein polymers with the minus (–) end embedded in spindle poles and the plus (+) end growing outward from the poles (see Figure 2). The α and β tubulin are the major structural component of microtubules and, together with microtubule-associated proteins (MAPs) and microtubule motors, make up the spindle (Dutcher, 2001; Nogales and Wang, 2006; Wittmann et al., 2001). In all eukaryotic cells, the mitotic spindle aligns all chromosomes on the metaphase plate and promotes the faithful segregation of the sister chromatids during cell division (Walczak and Heald, 2008). The mitotic spindle contains three kinds of microtubules: astral microtubules, kinetochore microtubules, and polar microtubules (see Figure 2). The plus end of the astral microtubules interacts with the cell cortex. This interaction performs crucial function of orienting the mitotic spindle with the axis of cell division (Karsenti and Vernos, 2001). Kinetochore microtubules, present in mitosis, attach the sister chromatids via kinetochores. After attaching to the kinetochore, the microtubules pull the sister chromatids towards opposing poles. Polar microtubules, whose plus ends interact with other microtubules from the opposite pole, are the third class of mitotic spindle. In general, they have the following three functions: During mitosis, polar microtubules maintain the structure of the spindle. In prophase, polar microtubules are responsible for pushing the centrosomes apart. Whereas, in anaphase B, polar microtubules push the spindle poles apart (summarized in (Lodish et al., 2007)).

Microtubule dynamics increases dramatically in mitosis (Howard and Hyman, 2003). The function of the mitotic spindle depends on rapid turnover of tubulin (Mitchison et al., 1986), which consist of the microtubules and the ability of kinetochores to keep attached to growing or disassembling microtubules (Rieder and Salmon, 1998).

The half-life of microtubules in a mitotic spindle is about 15 seconds, as compared to an interphase cell, where the half-life is about 5 minutes (Lodish et al., 2007). Microtubule dynamics can be affected by microtubule binding proteins. For example members of kinesin-13 family act as microtubule-depolymerizing factors (Desai et al., 1999) while the cytoplasmic linker protein (CLIP)-associated proteins (CLASPs), act as microtubule-stabilizing factors (Mimori-Kiyosue et al., 2005).

Additionally, microtubule-dependent motor proteins also contribute to spindle function. Kinesin-related proteins, such as Eg5, move toward the plus end of microtubules (Sawin et al., 1992) , whereas proteins like dynein moves toward the minus end (Steuer et al., 1990). These proteins create forces that either push the poles of the mitotic spindle further apart or pull the sister chromatids towards the spindle poles (Steuer et al., 1990). Any defect in this process will lead to potentially lethal errors in chromosome segregation.

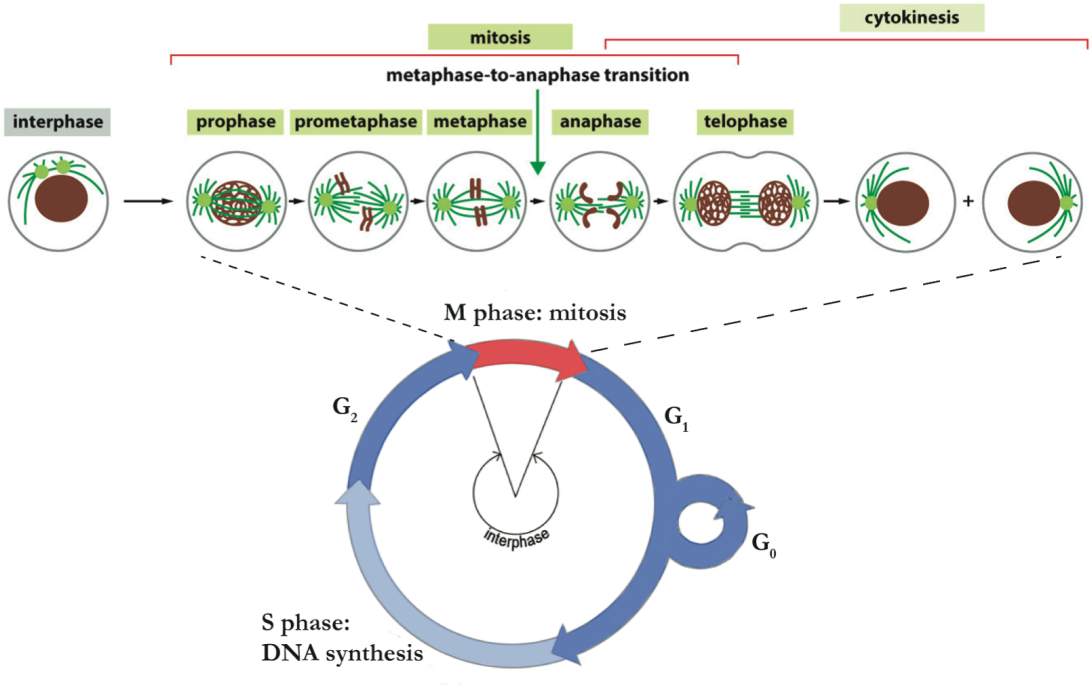

Figure 1 Eukaryotic cell cycle

The cell cycle can be divided into four phases. M phase, G_1 phase, S phase and G_2 phase. The major events happen in M phase and S phase, whereas G_1 and G_2 are the gap periods between the two phases. The G_0 is a resting state, in which some cells will resume proliferation and in which other cells will remain permanently. As we showed in the figure, mitosis includes five stages. From the metaphase to anaphase, there is a significant change in the biochemical state of the cell. The cytokinesis is the stage that the cells go back to interphase again. Adapt from (Alberts et al., 2014).

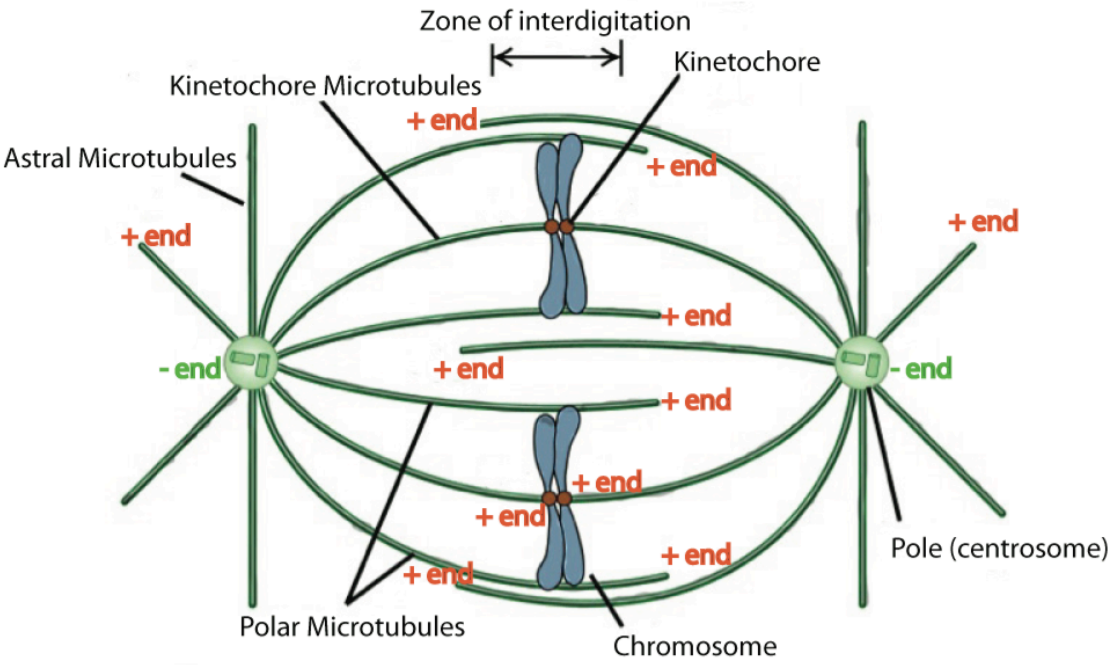

Figure 2 Schematic diagrams of Mitotic Spindly microtubules

The mitotic spindle contains three classes of microtubules (MTs). All the (–) ends of microtubules attach to the poles. Astral microtubules are linked to the cell cortex. Kinetochore microtubules attached to chromosomes through kinetochore. Polar microtubules overlap with each other with their distal (+) ends. Adapt from (Lodish et al., 2007).

1.3 The spindle assembly checkpoint

The spindle assembly checkpoint (SAC) is an evolutionary conserved safety mechanism, which delays the anaphase onset until each sister chromatid is attached properly to microtubules from opposite spindle poles (reviewed in (Foley and Kapoor, 2013; Lara-Gonzalez et al., 2012b)).

The SAC prevents the precocious sister chromatids segregation through the formation of the mitotic checkpoint complex (MCC) and subsequent inactivation of Cdc20, which is a cofactor of the E3 Ubiquitin ligase anaphase-promoting complex (APC/C) (see Figure 3). The APC/C is responsible for the progress of mitosis, according to the ubiquitylation and further proteasome-dependent degradation of securin and CyclinB1 (reviewed in (Lara-Gonzalez et al., 2012b)). When the Cdc20 becomes sequestered in the MCC complex, the APC/C cannot be functional, leading to a pause in mitosis. A single incorrectly attached or un-attached kinetochore will promote the generation of MCC, and can prevent the cell to be released to anaphase (see Figure 3).

If all the kinetochores attached to the microtubule properly, a set of additional proteins will help the cell to turn off the inhibition signal. One important protein is dynein, which helps to disassemble checkpoint proteins, like Mad1 and Mad2, from kinetochore (Funabiki and Wynne, 2013; Gassmann et al., 2010). The RZZ complex, implicated in targeting dynein/dynactin to kinetochore, is also involved in this inhibition process. Consequentially, Cdc20 will be released from MCC to activate the APC/C complex. APC/C poly-ubiquitinates securin and cyclin B1 continuously, causing their degradation by the proteasome. This leads to activation of Separase, which cleaves sister chromatid cohesion, leading to their separation to different sides of the cell. At the same

time, the degradation of cyclinB1 inactivates Cdk1, and promotes the cell to mitotic exit (see Figure 3).

Protein kinases and additional kinetochore-associated proteins are the core SAC components involved in this process, like Ser/Thr kinases Aurora B, budding uninhibited by benomyl 1 (Bub1), monopolar spindle protein 1 (MPS1), Mad1, Mad2, Bub3, Bub1-Related1 (BubR1). In higher eukaryotes, additional proteins are also involved in regulation the SAC activity, like p31^{comet}, Cdk1-CyclinB1, polo-like kinase-1 (plk1), Haspin, Nek2, CENP-E and dynein-associated proteins (summarized in (Musacchio and Salmon, 2007)).

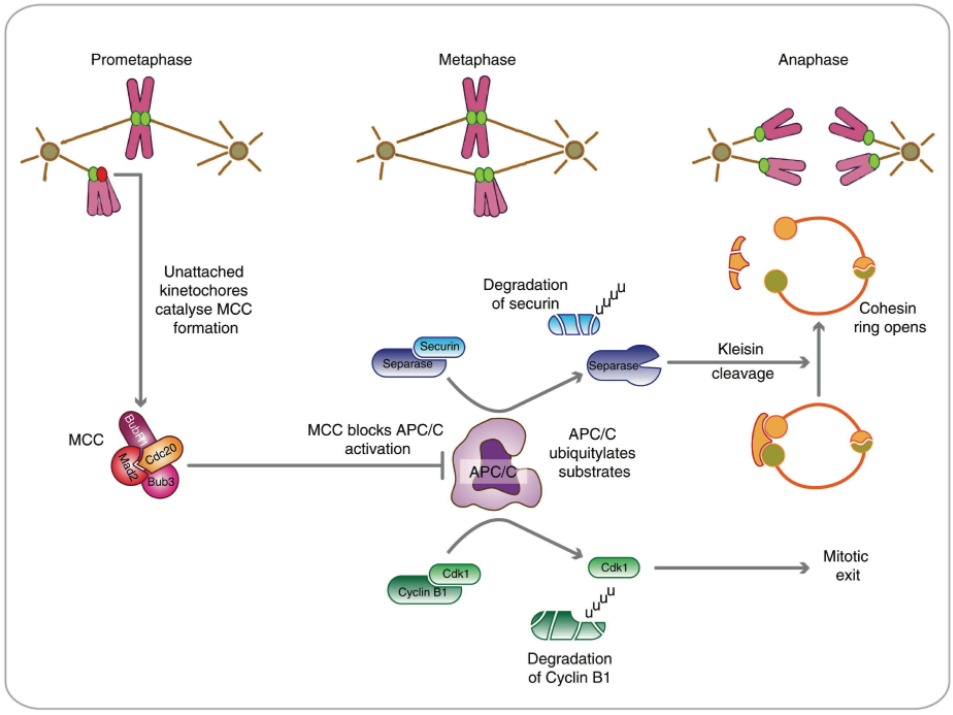

Figure 3 The process of spindle assembly checkpoint

At prometaphase, the unattached kinetochore, here shown in red, promotes mitotic checkpoint complex (MCC) formation, leading to the inhibition of APC/C. Once all the kinetochores attach to the spindle and are aligned in the right way, the formation of MCC stops and APC/C becomes active. The APC/C ubiquitylates both Cyclin B1 and Securin and the degradation of Securin leads to activation of Separase, which cleaves the Scc1 at the cohesion ring. Sister chromatids will separate, and the inactivation of Cdk1 will lead to mitotic exit. Adapted from (Lara-Gonzalez et al., 2012b).

1.4 The centromere and CENP-A

The simplest centromeres are found in fungi like Saccharomyces cerevisiae and are named point centromere (Pluta et al., 1995). They contain a ~150 base pair (bp) DNA sequence that is conserved among all chromosomes and that is also sufficient for kinetochore formation (Musacchio and Desai, 2017; Westermann et al., 2007). Kinetochores assembled on point centromeres, can only bind a single microtubule (Westermann et al., 2007).

Regional centromeres assemble kinetochores, which can bind to multiple microtubules (Cleveland et al., 2003a). They encompass kilobases of DNA and include both unique and repeated DNA elements (Pluta et al., 1995). For example, human centromeres, which belong to regional centromeres, consist of a large number of tandem repeats (Aldrup-Macdonald and Sullivan, 2014). But these repeats are not strictly required for centromere specification (Marshall et al., 2008; Musacchio and Desai, 2017). This strongly supports the idea that an epigenetic mechanism is required in the establishment and maintenance of the centromeric region (Allshire and Karpen, 2008; Musacchio and Desai, 2017).

In recent years, nucleosomes containing the histone H3 variant CENP-A have been defined as this crucial epigenetic factor (De Rop et al., 2012; Earnshaw, 2015; Stellfox et al., 2013), as the enrichment of CENP-A nucleosomes represents a marker of centromeres (Warburton et al., 1997).

CENP-A, which forms a dimer in CENP-A nucleosome (see Figure 4B), is highly divergent from histone H3 from primary sequences (schematically depicted in Figure 4A). The N-terminal tails of CENP-A and H3 are very divergent from each other: In CENP-A, the N terminus is enriched in arginine, whereas, the N-terminal of histone H3 is well

characterized with lysine. Therefore, the CENP-A tail cannot share the covalent epigenetic modifications of histone H3 (Panchenko et al., 2011). But The N-terminal of CENP-A has several post-translational modification (PTM) target sites, including the phosphorylation sites Ser7, Ser16, Ser18 (Panchenko et al., 2011; Zeitlin et al., 2001). All of this unique modification could be important for the centromere propagation and function (Cardinale et al., 2009; Nakano et al., 2008; Sullivan and Karpen, 2004).

In humans, the CENP-A targeting domain (CATD) region in the loop1 and alapha-2 helix region contains unique residues for CENP-A deposition. Similar domains to the human CATD have been identified as required for centromeric localization of CENP-A in *Drosophila* (Vermaak et al., 2002). It is not very clear how the CATD can determine centromeric localization of CENP-A. It is possible that the CATD influences stability, and/or specificity, of the interaction with DNA. This idea is supported by mutation experiments performed within loop I in CENP-A, which contacts with DNA extensively (Vermaak et al., 2002). Nearly all the mutation of conserved residues in Loop I lead to abolished targeting centromeres (Vermaak et al., 2002).

CENP-A is not only an epigenetic mark, but also has an essential function in kinetochore assembly (Musacchio and Desai, 2017; Stellfox et al., 2013; Weir et al., 2016). The kinetochore protein, CENP-C, interact directly with CENP-A (Guse et al., 2011). CENP-C is recruited to centromeres via its binding to the C terminal of CENP-A (Guse et al., 2011; Kato et al., 2013). However, the C-terminal tail of CENP-A is poorly conserved and is more hydrophobic than that of H3 (Leu-Glu-Glu-Gly-Leu-Gly and Glu-Arg-Ala in human CENP-A to H3) (Kato et al., 2013). Whereas, the higher hydrophobicity (not the specific amino acid sequence) of the C-terminal of CENP-A could be the key factor helping CENP-C to recognize and bind to CENP-A (Kato et al., 2013).

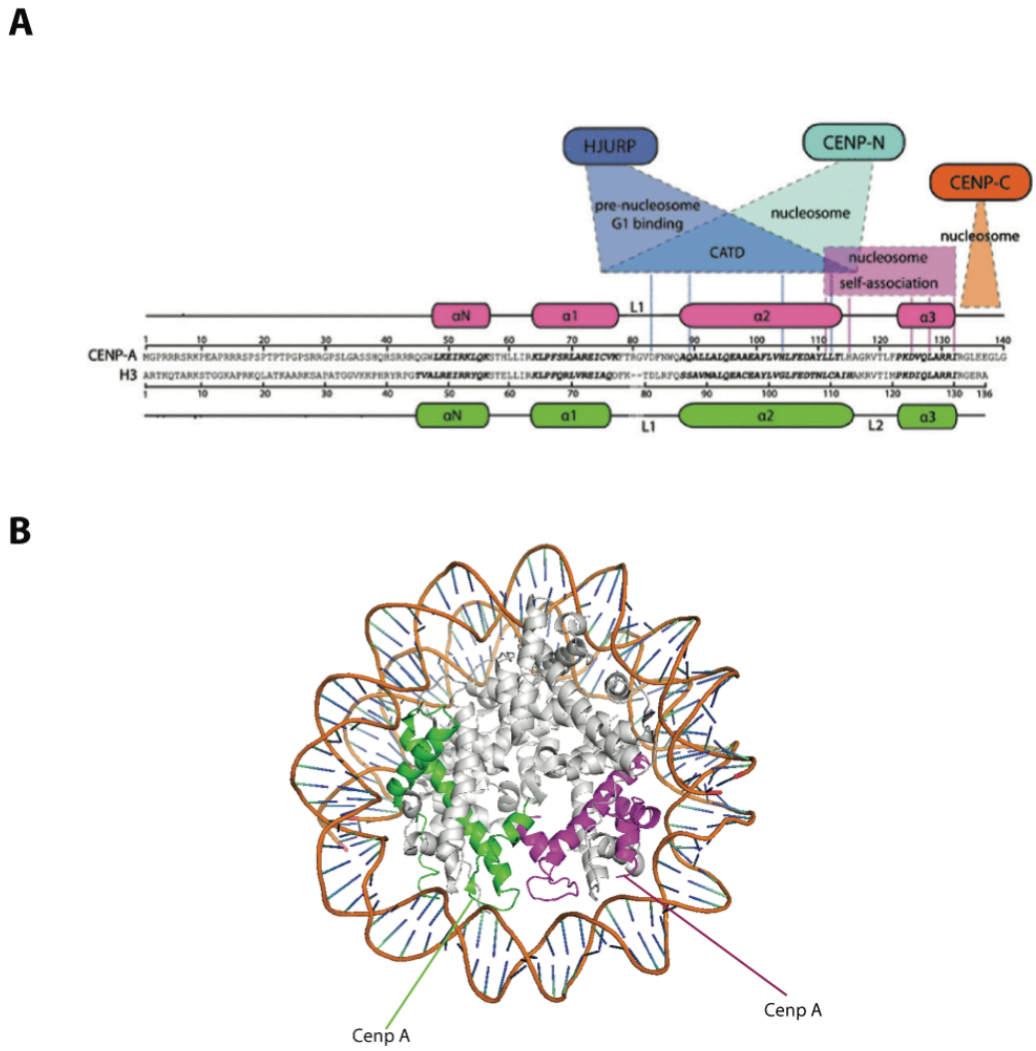

Figure 4 CENP-A nucleosome

(**A**) Primary sequence comparison of human CENP-A and H3.1. Most of the binding sites were labeled on the image. (**B**) Crystal structure of the CENP-A nucleosome (PDB ID: 3AN2). Two CENP-A molecules dimerize with each other in the nucleosome, which we shown in green and magenta. Adapted from (Stellfox et al., 2013).

1.5 Organization and function of kinetochores

For microtubules binding to the centromeric region of chromosomes, protein molecular machines localize to the centromere to build the kinetochores, which are required to act as a bridge between microtubules and chromosome (see Figure 5A-B).

Depending on the species, each kinetochore is able to attach to a different number of microtubules. *Saccharomyces cerecisiae* has the simplest system: one kinetochore binds to a single microtubule (Westermann et al., 2007). However, *Drosophila* and humans kinetochores bind to several microtubules (Musacchio and Desai, 2017; Sacristan and Kops, 2015). Even though the number of microtubules the kinetochore can bind differs between species, the overall structure and function of the kinetochore is conserved from yeast to *Drosophila* and humans (Cheeseman and Desai, 2008a; Kitagawa and Hieter, 2001; Musacchio and Desai, 2017; Santaguida and Musacchio, 2009).

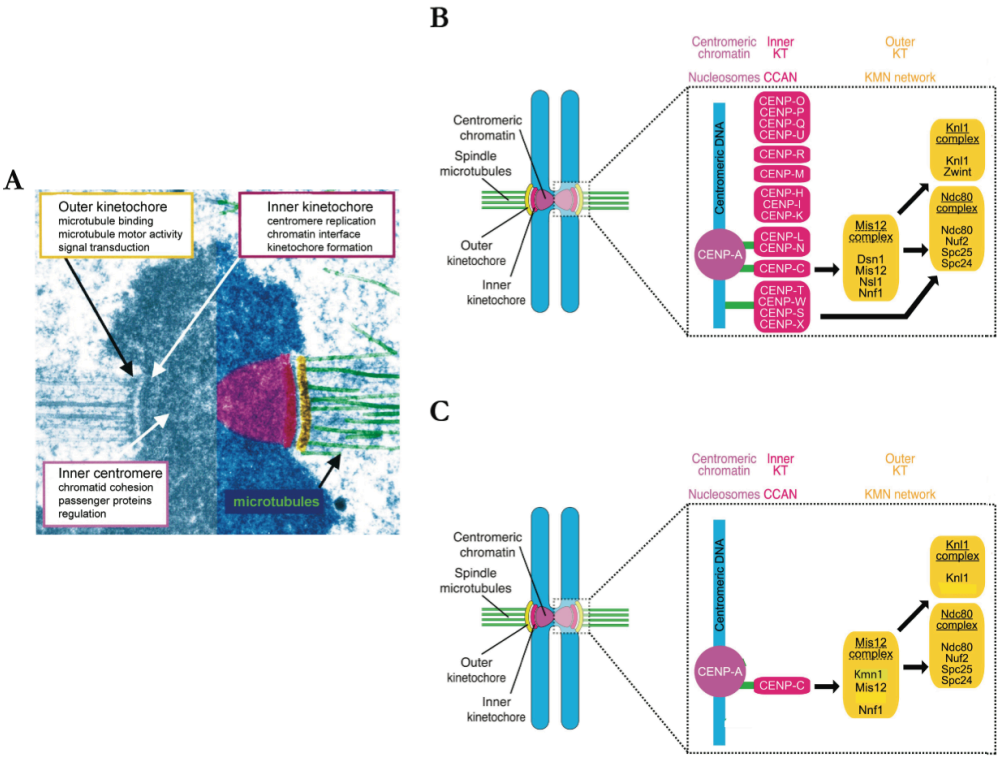

Figure 5 Overview of kinetochore organization

(**A**) Overall organization of the centromere. A mitotic chromosome has been sectioned along the plane of the spindle axis. (Red) The inner kinetochore, a region of distinctive chromatin composition attached to the primary constriction. (Yellow) The outer kinetochore is comprised of a diverse group of microtubule motor proteins, regulatory kinases, microtubule binding proteins, and mitotic checkpoint proteins; (**B**) The kinetochore has a trilaminar structure: 1. The constitutive centromere associated network (CCAN) is responsible for chromosome binding. 2. KMN (Knl1 complex/Mis12 complex/Ndc80 complex) network is the microtubule binding machinery that makes up the outer kinetochore. 3. The translucent layer in the middle; (**C**) Kinetochore in *Drosophila*. Adapted from (Cleveland et al., 2003b).

The primary function of the kinetochore is to ensure correct attachment of microtubules from opposing poles to each sister chromatid. To this aim, kinetochores perform four essential tasks: (i) they contribute a strong interface with centromere; (ii) they provide microtubule attachment sites; (iii) they ensure a feedback control mechanism to supervise microtubule attachments, the spindle assembly checkpoint (SAC); and finally (iv), they contain machinery that distinguishes correct from incorrect attachments to microtubules. Kinetochores contain over 100 individual proteins

(Cheeseman and Desai, 2008a; Santaguida and Musacchio, 2009). With the help of fluorescence microscopy, the location of most kinetochore proteins along the kinetochore axis has been determined at nanometer resolution (Joglekar et al., 2009; Schittenhelm et al., 2007). It is only now that we have been able to start to answer the question of how do the kinetochore proteins organize and interact with each other. Electron microscopy has shown that the vertebrate kinetochore has a trilaminar structure (Brinkley and Stubblefield, 1966; Rieder, 1982). As shown in figure 5A, it consists of an electron dense inner and outer kinetochore layers separated by an electron lucent middle layer. Figure 5B shows, different proteins can be found in each of the layers depending on their function within the kinetochore.

1.5.1 The Inner Kinetochore

As showed in the last section, the inner layer of the kinetochore contains the constitutive centromere-associated network (CCAN). All the CCAN subunits begin with CENP, for centromeric protein, followed by an alphabetic letter. CENP-A, CENP-B, and CENP-C were the first identified CCAN subunits from human centromere with anti-centromere antibodies (Earnshaw, 2015; Earnshaw and Rothfield, 1985). In vertebrates, CCAN contains more than 16 proteins (Amano et al., 2009; Ando et al., 2002; Earnshaw, 2015; Foltz et al., 2006; Hori et al., 2008; Izuta et al., 2006; Musacchio and Desai, 2017; Okada et al., 2006; Santaguida and Musacchio, 2009).

The CCAN subunits can be separated in the following sub-complexes (see Figure 4): the CENP-LN complex, CENP-HIKM complex, CENP-OPQRU complex, and the CENP-TWSX complex (Musacchio and Desai, 2017). According to bioinformatics

analyses, CCAN orthologs are also found in other organisms, like yeast (Asbury et al., 2011; Meraldi et al., 2006a).

CENP-C was one of the first CCAN subunits to be identified through auto-antibodies in serum of patients with immune dysfunctions (Earnshaw and Rothfield, 1985). Later, by antibody microinjection, CENP-C was identified to be required for kinetochore assembly (Tomkiel et al., 1994). It has been shown that CENP-C is the only conserved subunit of CCAN in the eukaryotes (Heeger et al., 2005b; Holland et al., 2005; Oegema et al., 2001b). The depletion study of CENP-C had the same phenotype as CENP-A disruption, including chromosome alignment and kinetochore assembly defects (Fukagawa and Brown, 1997; Kwon et al., 2007). In humans CENP-C broadly affected the recruitment of other CCAN subunits (Milks et al., 2009; Tanaka et al., 2009), Mis12 complex (Kwon et al., 2007; Liu et al., 2006; Milks et al., 2009), Ndc80 complex (Kwon et al., 2007; Liu et al., 2006), Mad1 (Liu et al., 2006), Mad2 (Kwon et al., 2007; Liu et al., 2006; Milks et al., 2009), Bub1 (Liu et al., 2006), BubR1 (Kwon et al., 2007).

CENP-C in humans has 943 amino acids (see Figure 6B). Most of the CENP-C sequence has been predicted as highly positively charged and unstructured (Screpanti et al., 2011a). CENP-C is the one of the two mechanisms that link inner and out kinetochore (Gascoigne and Cheeseman, 2011; Musacchio and Desai, 2017; Screpanti et al., 2011a). The N terminal region of CENP-C contains a Mis12C binding motif, and the first 20 amino acids can bind to Mis12C efficiently (See Figure 6B) (Screpanti et al., 2011a). Close to the N terminus there is a PEST (proline, glutamic acid, serine and threonine) rich region, which is responsible for binding to the CENP-H/I/K/M and CENP-N/L complexes (Klare et al., 2015b; Nagpal et al., 2015a). In the middle and close to the C terminal part, there are two CENP-A binding domain, which are both responsible for interaction with CENP-A nucleosomes (Kato et al., 2013). The central binding domain can promote both CENP-A nucleosome binding and kinetochore

targeting (Carroll et al., 2010a; Kato et al., 2013; Tanaka et al., 2009). The Cupin domain at the C terminus is responsible for dimerization of the protein, which may be important for formation of the kinetochore (Cohen et al., 2008; Sugimoto et al., 1997a; Yang et al., 1996). These essential roles of CENP-C are conserved in different organisms (Klare et al., 2015b).

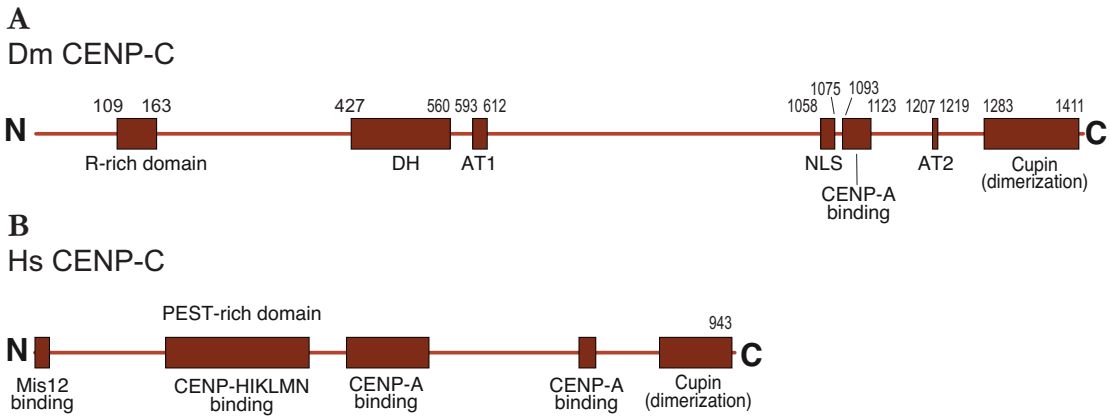

Figure 6 Schematic comparison of CENP-C in *Drosophila* and humans

(**A**) The functional Domains in the *Drosophila*: R-rich, arginine-rich; DH, Drosophilid CENP-C homologs; AT1 and AT2, AT hooks; NLS, nuclear localization signal; CENP-A binding motif is the CENPC motif; Cupin, a conserved dimerization domain at the C-terminal region (Heeger et al., 2005b); (**B**) CENP-C in humans has an N-terminal Mis12 binding domain (Hornung et al., 2014a; Przewloka et al., 2011a; Screpanti et al., 2011b), a domain for binding to the CENP-HIKM and CENP-NL complexes (Klare et al., 2015a; Nagpal et al., 2015b), and domains similar to these in *Drosophila*.

1.5.2 The outer Kinetochore

The outer kinetochore consists of Knl1 complex, Mis12 complex and Ndc80 complex (KMN network) (see Figure 7). The major function of the outer kinetochore is

that it offers the main platform for end-on microtubule binding and transduces the force generated by depolymerizing microtubules to move chromosomes (Cheeseman and Desai, 2008a; Musacchio and Desai, 2017). It is also the module that recruits the spindle assembly checkpoint proteins (Cheeseman and Desai, 2008b; Lara-Gonzalez et al., 2012a; Musacchio and Desai, 2017).

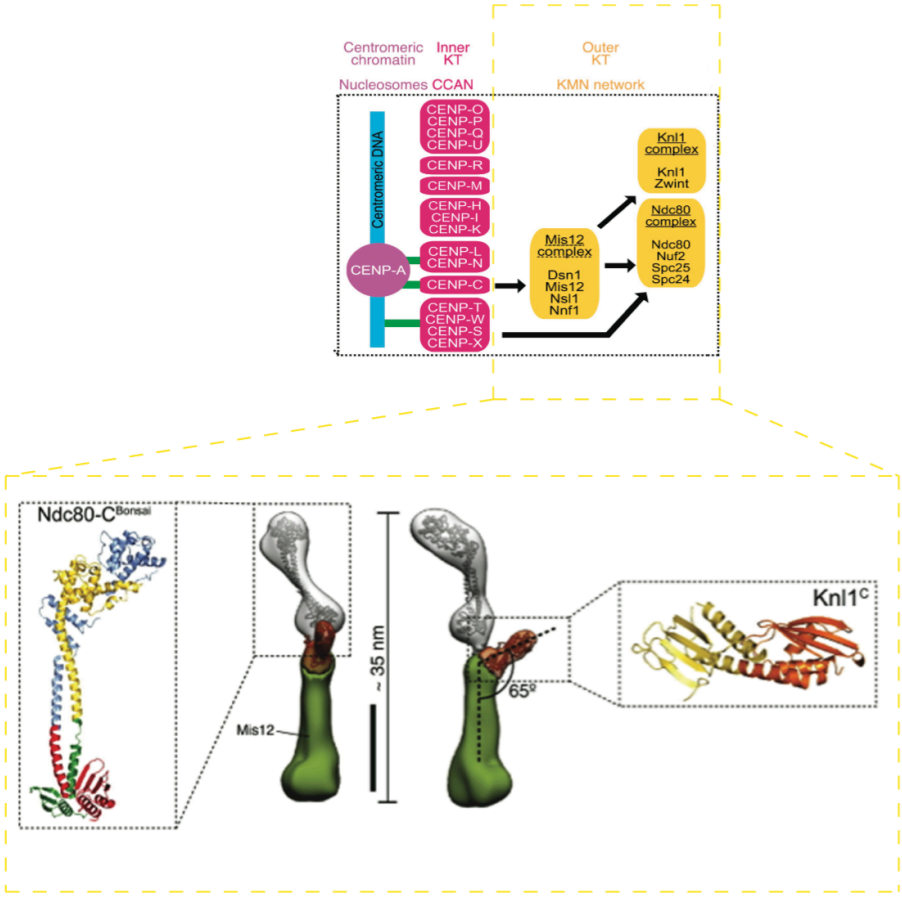

Figure 7 3D EM structure of Mis12-Ndc80Bonsai-Knl12106-2311 Complex

The 3D reconstruction used particles from negatively stained samples. We know the crystal structure of Ndc80-C^{Bonsai} (PDB ID: 2VE7) and $Knl1^{C}$ (PDB ID: 4NFA). The high-resolution structures were fitted into the density. Adapted from (Petrovic et al., 2014a).

The 4-subunit Ndc80 complex comprises the two sub-complexes Ndc80/Nuf2 and Spc24/Spc25, it is rich in coiled coil and highly elongated (about 57 nm at the long

axis) (Cheeseman et al., 2006; Cheeseman et al., 2004; DeLuca et al., 2006). It adopts a dumbbell-like structure whereby the two sub-complexes occupy at both ends of the dumbbell, the middle of which is coiled-coil region (Ciferri et al., 2005; Wei et al., 2005). The N-terminal regions of the Ndc80 and Nuf2 subunits mediate microtubule binding, and the C-terminal regions of the Spc24 and Spc25 subunits are responsible for the kinetochore targeting (Ciferri et al., 2005; Wei et al., 2006; Wei et al., 2005).

At the N terminal region of Ndc80 and Nuf2, which is responsible for microtubule binding, there is a pair of calponin-homology (CH) domain (Ciferri et al., 2008; Valverde et al., 2016; Wei et al., 2007). Cryo-electron microscopy studies for Ndc80 binding to microtubules indicate that there are direct interactions of the CH domain to the microtubule lattice with a spacing of 4 nm along each protofilament, which indicates that the CH domain interacts with both tubulin monomers (Alushin et al., 2012; Alushin et al., 2010).

Not only the CH domain, but also an N terminal tail of Ndc80, about 80 residues, has been showed to be involved in mediation of the binding with microtubules (Ciferri et al., 2008; DeLuca et al., 2006; DeLuca et al., 2011; Guimaraes et al., 2008). Up to nine sites in the human Ndc80 N-terminal tail can be phosphorylated by Aurora B kinase (Alushin et al., 2012; Alushin et al., 2010; Ciferri et al., 2008; DeLuca et al., 2006; Wei et al., 2007). This phosphorylation decrease the affinity between Ndc80 complex and microtubules (Cheeseman et al., 2006; DeLuca et al., 2006).

The other end of Ndc80 complex contains the Spc24 and Spc25, which target Ndc80 complex to the kinetochores (see Figure 7) (Ciferri et al., 2008; DeLuca et al., 2006; Petrovic et al., 2014a). The Spc24 and Spc25 subunits are structurally related and consist of N terminal coiled coil motifs followed by C terminal RING finger, WD repeat, DEAD-like helicases (RWD) domains (Ciferri et al., 2008; Malvezzi et al., 2013; Wei et al., 2006). These domains mediate kinetochore recruitment through interactions with at

least two distinct sequence motifs on the Nsl1 and Dsn1 subunits of Mis12 complex (Malvezzi et al., 2013; Petrovic et al., 2010b).

The Knl1 complex includes Knl1 and Zwint. The human Knl1 is 2316 amino acids. Except for the C-terminal 500 residues, most of Knl1 is disordered. The N terminal region of Knl1 contains a set of protein docking motifs, like a PP1 phosphatase binding site, which is very close to a microtubule binding site and has been implicated in checkpoint silencing (Espeut et al., 2012; Liu et al., 2010). In addition to sites for PP1 docking, there are also several Met-Glu-Leu-Thr (MELT) repeats, which are responsible for the recruitment of checkpoint components, including Bub1, BubR1, and Bub3 after phosphorylation by Mps1 kinase (Kiyomitsu et al., 2007; Krenn et al., 2012; Primorac et al., 2013; Yamagishi et al., 2012).

The C terminal end of Knl1 contains a predicted coiled-coil region, which is responsible for Zwint binding (The other knl1 complex subunit)(Kiyomitsu et al., 2011; Nekrasov et al., 2003; Pagliuca et al., 2009). The latter is also important in SAC activation mediated by RZZ recruitment (Rod, Zwilch and Zw10)(Kops et al., 2005; Starr et al., 2000; Wang et al., 2004). The C-terminal region of Knl1 proteins, containing tandem RWD folds, mediates binding to the Mis12 complex (Petrovic et al., 2014a; Petrovic et al., 2010b). Depletion of Knl1 in humans results in defects in the formation of kinetochore microtubule attachments and displaces Bub1 and BubR1 from kinetochores (Cheeseman et al., 2008; Kiyomitsu et al., 2007).

Mis12 Complex, which includes Mis12, Nnf1, Nsl1 and Dsn1, acts as an interaction hub that promotes KMN network assembly (Petrovic et al., 2014a; Petrovic et al., 2010b). Previous research in yeast and humans indicate that the Mis12 complex can be separated into two dimers, consisting of Dsn1-Nsl1 and Mis12-Nnf1 sub-complexes (Hornung et al., 2011a; Maskell et al., 2010a; Petrovic et al., 2010b). Depletion of Mis12 complex affects the recruitment of multiple proteins to kinetochore (Cheeseman et al.,

2004; Kline et al., 2006; Obuse et al., 2005). Negative stain EM revealed an elongated, approximately cylindrical structure, with one end bigger than the other, and with a long axis of around 22 nm (Hornung et al., 2011a; Maskell et al., 2010a; Petrovic et al., 2014a; Petrovic et al., 2010b). At one end, it binds both Ndc80 (Spc24 and Spc25 subunits) and the Knl1 complexes. Both Knl1 and Ndc80 complexes, mediate the interactions with Mis12 through RWD domains (See Figure 7). So Mis12 complex acts as an interaction hub for the RWD-containing proteins, and previous results in our lab indicate that the Nsl1 subunits is involved in binding with both Ndc80 complex and Knl1 complex (Petrovic et al., 2014a; Petrovic et al., 2010b)

As already discussed above, Mis12C also interact with the very conserved N0-terminal region of the CCAN subunit CENP-C, which acts as a linker between the inner and outer kinetochores (See Figure 5B) (Przewloka et al., 2011b; Schittenhelm et al., 2007; Screpanti et al., 2011a). Furthermore, Aurora B mediates this interaction by phosphorylating Ser100 and Ser109 on Dsn1 in Mis12C. This phosphorylation of Dsn1 promotes the binding between Mis12C and CENP-C (Kim and Yu, 2015; Rago et al., 2015a; Welburn et al., 2010). In *S.cerevisiae*, this positive regulation by Aurora B is conserved (Akiyoshi et al., 2013).

CENP-T provides the second linker between the inner kinetochore and outer KMN network with its unstructured N terminal region binding to the RWD domain in Spc24 and Spc25 subunits (See Figure 5B) (Altunkaya et al., 2016; Hori et al., 2013; Malvezzi et al., 2013). In addition, recent studies indicate that not only Ndc80C, but also Mis12C contribute to this second linkage, interacting directly with CENP-T (Huis In 't Veld et al., 2016; Rago et al., 2015b).

1.5.3 Specifics of *Drosophila* Kinetochore

Kinetochores from humans and budding yeast were often used as model, and large body of data has accumulated over the years pertaining to the structural and functional organization of kinetochores derived from these two organisms. These two kinetochores are quite similar, and only have limited differences, such as the absence of CENP-M in yeast (reviewed in (Musacchio and Desai, 2017)). During the evolution, the kinetochore had apparently changed many of its subunits. For example, in *Drosophila* and *C. elegans*, most CCAN subunits are missing; In *Bombyx mori*, most of the CCAN and also Nsl1 from Mis12C are missing (reviewed in (Drinnenberg et al., 2016)). However, the function of kinetochores in chromosome segregation is always conserved in different species. The analysis of special, structurally simplified kinetochore, will definitely help us to understand the key function of kinetochore.

1.5.3.1 The inner specific kinetochore

In certain organisms, including *Drosophila* and *Caenorhabditis elegans*, most of CCAN subunits have not been identified, suggesting that these kinetochores underwent significant structural simplifications in the course of evolution (Barth et al., 2014; Cheeseman and Desai, 2008a; Meraldi et al., 2006b; Przewloka et al., 2011a; Przewloka et al., 2007). The only residual CCAN subunit to be clearly recognizable in *Drosophila* is CENP-C (Heeger et al., 2005b; Moore and Roth, 2001; Oegema et al., 2001a; Saitoh et al., 1992). As we described before, CENP-C has been shown to act as a linker between CENP-A (CID in *Drosophila*) in the centromeric chromatin and the Mis12C in the outer

kinetochore, a function that appears to be conserved in most or all eukaryotes (Klare et al., 2015b).

Drosophila CENP-C contains 1411 amino acid, as we show in Figure 6A (Heeger et al., 2005a; Orr and Sunkel, 2011; Przewloka et al., 2011b). The Drosophilids CENP-C proteins show high similarity in some functional domains with human (see Figure 6A-B). The N terminal part of *Drosophila* CENP-C is sufficient to recruit core kinetochore component (Przewloka et al., 2011b). But the binding details were not clear when I started the project, and this would be a major point in this thesis. The most divergent sequence region is that in *Drosophila*, there is an arginine-rich region and another Drosophilids homology region (Heeger et al., 2005a). The equivalent position of Drosophilids homology region in humans is a second centromere localization domain (Song et al., 2002; Yang et al., 1996). The functions of these two regions of *Drosophila* CENP-C are not clear. There are also two extra AT hooks (AT1 and AT2), which mediate binding to the minor grove of DNA (Heeger et al., 2005a).

Despite its divergence, CENP-C in *Drosophila* is an essential factor for CENP-A^{CID} assembly at *Drosophila* centromeres and is required for centromere identity. MEI-S332 (*Drosophila* homolog of Shugoshin) and chromosomal passenger proteins fail to localize in CENP-C depleted cells (Erhardt et al., 2008; Orr and Sunkel, 2011). All of this indicates that in *Drosophila*, CENP-C plays an essential role in overall centromere and kinetochore organization, a role that might be shared with the CCAN protein complexes in other systems.

1.5.3.2 The outer specific kinetochore

Besides the loss of most CCAN subunits in the inner kinetochore, in *Drosophila* additional evolutionary changes affected the composition of the outer kinetochore, and

in particular of the Mis12C complex. These changes include the apparent loss of one homologue of the Mis12 complex, and the appearance of two paralogues of the Nnf1 (Nnf1a and Nnf1b). Previous studies indicated that *Drosophila* Mis12 complex lost the Dsn1 subunit (Przewloka et al., 2007; Schittenhelm et al., 2007). However, there are no strong biochemical or cell biology evidences to support this conclusion. In general, the Mis12 complex in *Drosophila* also contains four subunits: Mis12, Nnf1a, Nnf1b, and Kmn1, but Nnf1a and Nnf1b are closely related paralogs, and it was unclear if they can be incorporated in the same or in different complexes (Przewloka et al., 2007; Schittenhelm et al., 2007). In human cells, depletion of Mis12 results in loss of some kinetochore proteins without affecting centromeric localization of CENP-A, whereasdepletion of CENP-A partially or completely affects centromeric localization of Mis12 (Liu et al., 2006). In *Drosophila*, Mis12 localization depends on CENP-A^{CID} (Przewloka et al., 2007). The recruitments of Mis12 complex to centromeres in *Drosophila* is different than in humans. In cultured cells, Mis12 and Nnf1a subunits localize to centromeres in interphase and thus in advance of the mitotic localization of Kmn1 and Nnf1b (Venkei et al., 2011a). Even though different subunits of Mis12 complex in *Drosophila* localize at kinetochores at different time, the Mis12 complex still acts as a single functional unit during mitosis (Venkei et al., 2011a). Mutations of both *mis12* and *kmn1* genes showed similar developmental and mitotic defects in *Drosophila* (Venkei et al., 2011a). When I started this project, there was still no high-resolution structure of Mis12 complex. How do the Mis12 complex subunits in *Drosophila* interact with each other? This is the question on which I became engaged during my doctoral work and on which my thesis focuses.

Drosophila expresses homologs of the Ndc80 complex subunits (Przewloka et al., 2007; Schittenhelm et al., 2007). In both yeast and vertebrate Ndc80 complexes, Spc24 is 22-24 kDa (Cheeseman et al., 2006; Ciferri et al., 2005). The molecular weight of

Drosophila Spc24 is about 11 kDa, which might correspond to the C terminal globular part of Spc24 domain in vertebrates (Schittenhelm et al., 2007). The globular domains of Spc24 and Spc25 in humans are essential for kinetochore targeting, as they interact directly with the Mis12 complex and also CENP-T (Hori et al., 2013; Huis In 't Veld et al., 2016; Nishino et al., 2013; Petrovic et al., 2014a). In *Drosophila*, both Mis12 and Knl1 are required for Ndc80 localization at KT, unlike human where both Mis12 and CENP-T are required (Schittenhelm et al., 2007; Venkei et al., 2012b). The C terminus of DmSpc105R (Spc105-related, homologous to human Knl1) lies at the junction of the DmMis12 and DmNdc80 complexes on stretched kinetochores (Venkei et al., 2012b),

The Zwint subunit in the Knl1/Spc105R complex is also apparently absent in *Drosophila* (Przewloka et al., 2009; Przewloka et al., 2007; Schittenhelm et al., 2009b; Schittenhelm et al., 2007; Venkei et al., 2011b; Venkei et al., 2012a). Knl1/Spc105R in *Drosophila* contains 1959 amino acids and results of Y2H (yeast 2-hybrid) assays demonstrated that the N terminus of Knl1/Spc105R also binds to Bub1 (Schittenhelm et al., 2009a), similarly to Knl1/Spc105R in humans (Kiyomitsu et al., 2007). The middle part of Knl1/Spc105R contains the repeat region, which amounts to 44% of the full-length *Drosophila* sequences. However, in vivo experiments with the equivalent region of human Knl1 had not shown any essential function of this part . The C-terminal region of Knl1/Spc105R interacts with Mis12 complex in vivo , similar with C terminal of Knl1/Spc105R in humans (Petrovic et al., 2014a; Schittenhelm et al., 2007).

1.6 The Objectives

Our work on the outer kinetochore of *Drosophila* was motivated by its considerable simplification in comparison with its counterparts in other organisms, like in humans and yeast. In this thesis, I have used biochemical reconstitution and biophysical characterization as an entry point to characterize the outer kinetochore of *Drosophila* and its interaction with the sole remaining CCAN component CENP-C. This work will shed light on the structure and function of the simplified outer kinetochore in *Drosophila*. It helps us to understand the functions of the core out kinetochore subunits, which should be conserved during the evolution.

2 MATERIALS and METHODS

2.1 The cloning of the DNA constructs

All constructs were sub-cloned into bacterial expression vectors. The affinity tags were placed at the N or C terminal end to aid purification. All the genes were amplified by polymerase chain reaction (PCR) using Pfu polymerase. All the DNA restriction endonucleases were purchased from NEB and used as the manuals. All the clones produced in this study were amplified in DH5 α or Top10 chemically competent cells and stored in -80°C. The Mini Qiagen kits were used to perform mini preparations of plasmids. All the procedures were followed the manufactory protocols. All DNA constructs were conformed by sequencing and the service was from Beckman. All the chemicals that we used were from Sigma Aldrich, Hamburg, Germany.

The cDNA for DmSpc105$^{1707\text{-}1960}$ was amplified by the polymerase chain reaction (PCR) from the pOT2 vector containing the full-length DmSpc105 sequence (isoform A; a generous gift of Christian Lehner's Lab in University of Zurich) and subcloned into the fourth cassette of pST44 (see Figure 8) (Tan et al., 2005). DNA sequences optimized for expression in E.coli of DmMis12, DmNnf1a, DmNnf1b, DmKmn1 and full length DmCenp-C were obtained from GeneArt. DmCenp-C fragments were amplified by PCR and subcloned into the pETDuet-MBP8His vector, a modified version of the pETDuet vector (from Novagen) generated in house by Dr. Dongqing Pan. Sequences encoding variant versions of the DmMis12 complexes were generated in the pST44 system using standard restriction digest based cloning as per the manufacturers instructions. The QuikChange Mutagenesis kit (Agilent Technologies) was used to generate all genetic mutations in the plasmid DNA used in this study as the manufacturers instructions.

In order to express the protein complexes *in vitro*, most of the constructs (especially the complex) were sub-cloned to pST44 polycistronic system (Tan et al., 2005). The pST44 polycistronic system utilizes the concept of multiple individual translation cassettes. Each cassette contains a coding region defined by START and STOP codons. It also includes the translational initiation signals, such as the Shine-Dalgarno sequence and translational enhancers (see Figure 8). Combination of these features makes the vector sufficient for the *E. coli* translational machinery to start and sustain translation of the mRNA into the desired protein.

In the pST44 system, transfer or donor plasmid DNA for each translation cassettes is separate and independent. The coding region of each gene is first sub-cloned into the transfer plasmid. There are 72 transfer plasmids to choose from in the pST44 system suite. Each transfer plasmids contain modular affinity tags that contain cleavable or non-cleavable sites. Traditionally, four subunits can be co-expressed in the pST44 polycistronic expression system. It is also possible to reconstitute complexes containing 8,12, or even more proteins using the pST44 system (Tan et al., 2005). During the expression process, different individual subunits or sub-complexes may co-exist. We can use conventional chromatography methods to purify each subunit. Or it is also possible to use different affinity tag in different subunit, and then try to separate with affinity purification steps. In this thesis, I used both strategies.

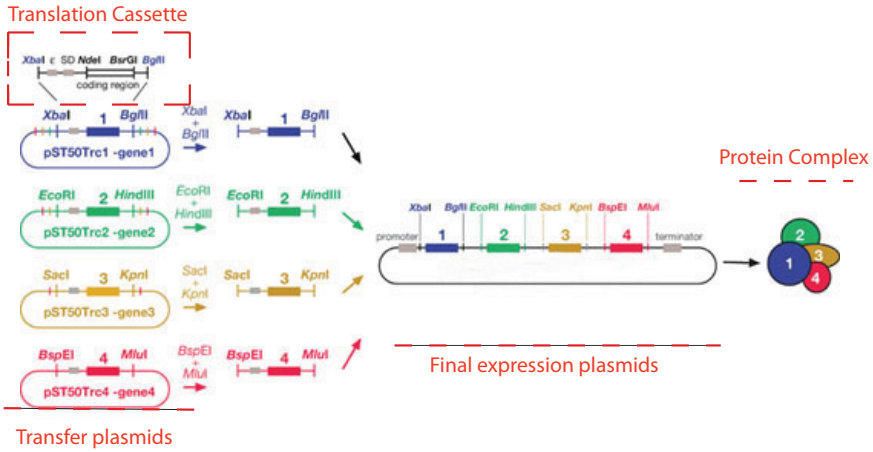

Figure 8 Schematic of the polycistronic expression systems

The pST50Trc1-4 plasmids were used to provide the translation cassettes. Unique restriction enzymes were used to subclone the translation cassettes into pST44. Translational enhancer (ε) and the Shine-Dalgarno sequence (SD) were included into every translation cassettes. Translation cassettes 1, 2, 3, 4, which include unique restriction sites and coding regions, are shown in blue, green, yellow, and red respectively. Adapt form (Selleck and Tan, 2008).

2.2 Protein expression and Purification

BL21 (DE3) Rosetta cells were used to express all recombinant proteins. Cells were grown in Terrific broth at 37°C to an OD_{600} of approximately 0.8. Protein expression was induced by addition of 0.1 mM IPTG at 20°C, and cells were incubated overnight. Cell pellets were re-suspended in binding buffer (20 mM Tris/HCl pH 8.0, 300 mM NaCl, 5% (v/v) glycerol, 1 mM EDTA, 1 mM TCEP) lysed by sonication and cleared by centrifugation at 10000x*g* for 30 min.

In order to carry out the all the biochemistry and biophysical experiments in this thesis, I needed to optimize all the purification procedures in order to obtain highly pure proteins in high concentrations. An efficient purification protocol contains the fewest number of steps and must minimize the sample handling to reduce the loss of protein. Most purification strategies contain four general steps. First, the target protein is isolated and stabilized. Second, the majority of impurities are removed. Third, a polishing step is applied if required for the removal of remaining impurities. Finally, the buffer is exchanged and the purified protein is concentrated. All of these steps exploit different properties of the target protein or protein complex, for example, protein size, surface charge, ligand specificity, hydrophobicity, or isoelectric point (PI). The following paragraphs will give a brief summary of the basic chromatographic techniques used in this study.

2.2.1 Affinity chromatography

Specific binding interaction between molecules is the base of affinity chromatography. A specific ligand is coupled to a surface (solid support), and the

mixture is passed over the solid support. The molecules, which have specific binding affinity to the ligand, will bind to the solid support. The other impurities will be washed away. The bound targets can be stripped from the solid support by a competing ligand or protease cleavage. The result is a more pure protein or protein complex.

Affinity purification involves three steps. First, raw samples are incubated with the affinity support for 2 hours. This will allow the target molecule in the sample to bind to the immobilized ligand. Secondly, the appropriate buffers, which can maintain the binding interaction and the protein activity, are used to wash away contaminants. Finally, the target protein is dissociated and recovered from the immobilized ligand.

In our study, most of the proteins were tagged with 8 consecutive histidine residues at the N terminus or C terminus of the proteins. The HisTrap chelating columns contains cross-linked agarose beads to which iminodiacetic acid has been coupled via a seven-atom spacer. The resin requires a suitable metal ion to activate its binding affinity; in this study, Ni^{2+} was used. Histrap Chelating columns (GE Healthcare), charged with suitable metals ions, will retain the target protein tagged with the flexible His tag. Target proteins were released from the matrix by addition of high concentrated imidazole, which competes with the protein's his tag to bind to the matrix.

2.2.2 Ion exchange chromatography

Because each protein has a different net surface charge, different proteins exhibit different degrees of interaction with charged chromatography media. Ion exchange chromatography exploits these differences. The charged amino acid groups on the protein, which contribute to the net surface charge, have different pKa values (acid ionization constant) depending on their structure and chemical microenvironment

including being highly pH dependent. The relationship between net surface charge and pH is unique for each specific protein. The interaction between charged proteins and oppositely charged ion exchange chromatography (IEX) media is based upon this principle. The overall charge status of the target protein can be adjusted by choosing the appropriate pH. When a protein is in a pH buffer equivalent to its isoelectric point (pI) it has no net charge and won't interact with a charged medium. In a buffer with a pH above its pI, a protein will bind to a positively charged medium. While, if in a buffer with a pH below its pI, the proteins will bind to negatively charged medium. Other factors can also affect the binding of proteins to IEX media, including van der Waals forces and nonpolar interactions, but in general these only account for a small part of the binding interaction.

Differential elution of species bound to the IEX column is achieved by altering the buffer. In general, the concentration of an ionic moiety competing for the binding is raised or the pH is changed. Strong or week matrices can be used in ion exchange. Strong ion exchange matrices are ionized across a wide range of pH levels. Weak matrices are charged within a narrower pH range.

2.2.3 Size exclusion chromatography

Size exclusion chromatography (SEC) separates proteins according to the difference in size as they pass through a SEC medium packed in a column. Unlike the other two purification methods, the proteins do not bind to the medium of SEC. The significant advantage of SEC is that the buffer conditions can be varied to suit any special requirements. Whether biomolecules are sensitive to pH changes, cofactors and harsh environmental condition or not, they can be purified by SEC. Purified proteins can

be pulled together in any chosen buffer condition. SEC media is a porous matrix of spherical particles that are chemically and physically stable. They also lack reactivity and adsorptive properties. Samples are eluted in isocratic flow, and there is no need to use different buffers during the separation. SEC can also be used to remove small molecules, desalting, or exchanging buffers.

A SEC column has a total volume (V_T) made up by three components: the volume of the external solvent V_0, the solid volume of the gel particles V_g and the internal volume of the porous particles V_i. The solute, which is totally excluded by the resin, will elute after a volume equal to the V_0. Conversely, a solute small enough to enter the porous beads will be slowed down and elute later. The sample volume ideally should be less than 2% of the V_T , in order to prevent loss of resolving resolution during the purification. Since the efficiency won't be affected by the sample concentration, small volumes of relatively high concentrated proteins are best.

2.3 Biophysical and structural methods

2.3.1 Analytical ultracentrifugation

Analytical ultracentrifugation (AUC) has broad applications for the study of macromolecules in a wide range of solvents. There are three optical systems available for the analytical ultracentrifuge: absorbance, fluorescence, and interference. The analytical ultracentrifuge permits precise and selective observation of sedimentation in real time. Unlike many commonly used methods, samples are characterized in their native state with biological solution conditions in analytical ultracentrifugation.

There are two types of analyses in the analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium. In sedimentation velocity, hydrodynamic theory is used to interpret the movement of solutes to define the size, shape and even interactions of macromolecules at high centrifugal fields (Howlett et al., 2006). Sedimentation equilibrium is a thermodynamic method. The equilibrium concentration gradients are analyzed to define molecule mass, assembly stoichiometry, association constants and solution non-ideality at low centrifugal fields (Howlett et al., 2006; Laue, 1995).

With the help of modern analysis software, we can use sedimentation velocity to determine the homogeneity of the sample. We can also define whether it undergoes concentration-dependent association reactions. According to more thorough model-dependent analysis of velocity and equilibrium experiments, we can get a picture of the nature of all the species in solution and their interaction relationship.

In this study, sedimentation velocity experiments were performed in an Optima XL-A analytical ultracentrifuge (Beckman Coulter, Palo Alto, US-CA) with Epon charcoal-filled double-sector quartz cells and an An-60 Ti rotor (Beckman Coulter, Palo

Alto, US-CA). Samples were dialyzed against buffer (20 mM Tris pH 8, 0.15 M NaCl and 1 mM TCEP) that was used as a reference. Samples were centrifuged at 42,000 rpm at 20 °C and 500 radial absorbance scans at 280 nm were collected with a time interval of 1 min. The data were analyzed using the SEDFIT software (Schuck, 2000) in terms of continuous distribution function of sedimentation coefficients (c(S)). The protein partial specific volume was estimated from the amino acid sequence using the program SEDNTERP. Data was plotted using the program GUSSI.

2.3.2 Negative stain electron microscopy

For many types of microscopy, improvement of the phase contrast is always a major target for the common user. In electron microscopy (EM), heavy metal salts, which are derived from molybdenum, uranium, or tungsten, are usually used for staining. Heavy ions can readily interact with the electron beam and produce phase contrast. The general steps of negative staining electron microscopy are as follows:

First, we prepared carbon-coated grids, and deposited a small drop of target protein complex (about 8 μl) on the carbon-coated grid. Secondly, the proteins were allowed to settle for about one minute, and were blotted dry on the grid. Finally, the grid was covered with a small drop of the stain (2% Uranyl Formate). After 30 seconds, this drop was also blotted dry, and the sample is ready of EM.

In negative stain microscopy, the electron beam primarily interacts with the stain, which had filled in around the protein samples. In a well-stained sample, the stain uniformly covers the sample and is excluded from the volume occupied by the sample. As the electron beam passed through the grid, it is deflected by its interactions with the sample and stain. The deflection through protein is less than through the stain rich

regions. The objective aperture, located below the sample, filters out highly deflected electrons. The size of the aperture is adjusted to determine the contrast and resolution of the final image from the screen. The stain absorbs electrons in much higher amounts, enhancing the contrast of the images. There is some drawbacks to negative stain EM: for example, the particle is distorted during the staining process, and the resolution is limited to approximately 20 Å under optimal conditions.

In our experiments, the Mis12 complex was diluted to 15 nM for EM grid preparation. The protein sample was adsorbed onto glow-discharged carbon-coated grids prior to negative staining with Uranyl Formate (SPI supplies/Structure Probe). Samples were imaged with a JEOL1400 microscope equipped with a LaB6 cathode operating at 120 kV. Images were recorded at low-dose conditions at a magnification of 67,200 on a 4k × 4k charge-coupled device camera (TVIPS GmbH).

2.4 Bio-analytical methods

2.4.1 Sodium dodecyl sulfate polyacrylamide gel electrophoresis

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) is used for separation of proteins majorly based on their molecular weight. Under the electrical field, the proteins are separated based on their differential rates of migration through a gel. The SDS, in the loading buffer, is used to destroy the tertiary structure and mask the intrinsic net charge of the protein. In combination with boiling, and addition of a reducing agent (e.g. DTT), the tertiary structures of proteins are disrupted in a very efficient way. The SDS binds uniformly to the linear proteins, and it coats the protein with a uniform negative charge. This means that the charge of the linear protein, coated with SDS, is now approximately proportional to its molecular weight.

The gels also contain SDS, which helps to make sure that the protein remains linearized and charges are masked throughout the run. The gel matrix used for SDS-PAGE is polyacrylamide from Sigma Aldrich. Gel matrices can be easily made up at different concentrations to produce different pore sizes. In our experiments, most of the gels are made in house, and contain 12% to 14% acrylamide.

All the samples were mixed with 5x SDS-PAGE loading buffer and denatured for 5 min at 96 °C. On each gel, a molecular weight marker (Precision Plus Protein Standard, Bio-Rad, Hercules, US-CA) was loaded in addition to the samples for comparison. The gels were placed into a Mini-PROTEAN Tetra Cell (from Bio-Rad) with running buffer and a voltage of 145 V was applied until the dye of the loading buffer exited the gel. Proteins were visualized by Coomassie staining and subsequent distaining with 10 % (v/v) acetic acid.

2.4.2 Protein concentration test

Protein concentration was determined by both Bradford assay and UV absorbance. The Bradford assay is a colorimetric assay for measuring total protein concentration. It is based on the Bradford dye-binding method (Bradford, 1976). In a standard procedure, the assay is used with protein samples, whose concentration is between 200 and 1,400 μg/ml. The method is based on the color change of Coomassie brilliant blue G-250 dye, which binds to primarily basic (especially arginine) and aromatic amino acid residues. The assay can be used for both proteins and polypeptides, with molecular weights more than 3 kD. For this purpose, 1-10 μg of protein was mixed with 1 mL of Bradford solution (Thermo Fisher Scientific, US-MA). After incubation for 3 min the absorbance at a wavelength of 595 nm was measured spectrometrically by comparison to a reference sample. Protein concentration was determined by use of a standard curve, which was made with known concentrations of BSA.

Determining the protein concentration in a solution is also possible by a simple spectrometer. The Tyrosine and Tryptophan are responsible for the absorption of radiation in the near UV (in our experiments, we used 280 nm). This method is quite simple, and the samples are even recoverable. However, the excitation of DNA in the 280 nm may be 10 times higher than that of protein. This implies that if the sample contains even a small fraction of DNA, the absorption is greatly influenced. The determination of protein concentrations via UV absorbance at a wavelength of 280 nm is based on the Beer-Lambert law. The measurement of absorbance was carried out with a NanoDrop spectrophotometer ND-1000 (peQLab, Erlangen, DE).

2.4.3 Western blotting

Western blots were used to probe specific proteins in cell lysates. The first step in the Western blotting procedure is to separate the protein complexes using SDS-PAGE. Instead of staining the run gel with Coomassie, it was transferred onto a second matrix, a polyvinylidene difluoride (PVDF) membrane (Bio-Rad Laboratories, Munich, DE), via electrophoresis in 25 mM Tris, 200 mM glycine and 20 % (v/v) methanol by applying 80 V for 90 min. Protein transfer was checked by staining with Ponceau S. The membrane was then blocked to prevent nonspecific binding of the antibodies to its surface. The membrane was blocked for 1 h at room temperature in PBS-T supplemented with 5 % (w/v) milk powder. The transferred membrane was incubated with the indicated primary antibodies, which had been diluted in PBS-T supplemented with 5 % (w/v) milk powder. This incubation was performed overnight at 4°C. After five washes, each for 5 min, with PBS-T, the membrane was incubated for 1 h at room temperature with HRP conjugated mouse anti-rabbit antibody (Amersham Biosciences, Piscataway, US-NJ) diluted in PBS-T supplemented with 5 % (w/v) milk powder. The membrane was washed five more times (5 min each) with PBS-T. A highly sensitive chemiluminescence detection reagent (from GE Healthcare Life Science) was added according to the manufacturer instructions to the membrane and the chemiluminescence signal was detected with a developer machine on film.

2.4.4 Analytical SEC migration shift assay

Analytical size-exclusion chromatography is the most common way to detect the binding of proteins in solution. As stated before, size exclusion chromatography (SEC)

separates molecules according to their hydrodynamic radius. Any changes in the hydrodynamic radius can can be used as an indicator of binding between respective proteins in solution.

Analytical size-exclusion chromatography experiments were performed on calibrated Superdex 200 5/150 column (GE Healthcare). The respective proteins were mixed at the concentrations of 10-20 µM in a final volume of 50 µl. Samples containing single proteins in size exclusion chromatography buffer (20 mM Tris, 150 mM NaCl, 1 mM TCEP) served as controls. In order to detect complex formation, proteins were typically mixed at 1:1 (molar ratios) and incubated for 2 hours on ice before injection onto the size exclusion chromatography column. The columns were equilibrated in size exclusion chromatography buffer and all samples were eluted under isocratic conditions at 4°C in size-exclusion chromatography buffer at a flow rate of 0.2 ml/min. Elution of proteins was monitored at 280 nm. SDS-PAGE, followed by Coomassie staining, was used to detect proteins. Subsequently, the chromatograms and the corresponding SDS-PAGE gels of the samples were compared for the single proteins and the protein mixtures.

2.4.5 Limited proteolysis

For limited proteolysis on MAK and MBK, the proteins were subjected to cleavage by four proteases with different cleavage sites: Trypsin (cleaves Arg and Lys), Chymotrypsin (cleaves large hydrophobic amino acids: Trp, Tyr, Phe. Also Leu and Met while long incubations), Papain (cleaves basic amino acids, Arg or Lys) and Subtilisin (cleaves large uncharged residues). Protease stocks at 1 mg/ml were diluted 1:10, 1:100

and 1:1000 in protease dilution buffer (20 mM HEPES pH 7.5, 50 mM NaCl, 10 mM $MgSO_4$). Proteases were stored at -80°, whereas the dilutions stored at -20°C.

6 µg of protein or protein complex was incubated with 3 µl of diluted protease, in 10 µl final reaction volume for 30 minutes at 4°C. Protein without any protease was used as a negative control, while the tube with the highest protease concentration served as a migration control. 5 µl of 5×SDS loading buffer was added to the samples in order to stop the reaction. The samples were then boiled at 95°C for 5 minutes and visualized on a Coomassie stained SDS-PAGE gel. In order to obtain reproducible results, aliquots of the same protease in all of the aforementioned dilutions were flash frozen in liquid nitrogen and stored at -80°C.

In order to combine limited proteolysis with electrospray ionization mass spectrometry, the small-scale reactions needed to scaled-up. We chose the most effective protease, with suitable dilution for scale-up. The SEC column was first equilibrated with buffer (20 mM Tris, 150 mM NaCl, 1 mM TCEP). The protein complexes were incubated with the chosen protease for 30 minutes at a suitable dilution identified during small scale testing. When the incubation was completed, 100 mM AEBSF was added to stop the reaction, and then the samples were directly loaded onto the SEC column. Fractions from the SEC were run on a SDS-PAGE gel and then stained with coomassie to identify the best fractions for electrospray ionization mass spectrometry

2.4.6 Electrospray ionization mass spectrometry

The Electrospray ionization (ESI) mass spectrometry (MS) was used to determine the molecular weight of molecules. The ESI source is used at atmospheric pressure. The protein solution was sprayed from a small tube into a strong electric field, which also

contained a flow of warm nitrogen to assist desolvation, increasing the charge on the evaporated droplets. The multiply charged ions enter the analyzer.

The proteins can be ionized without denaturation. For the analysis of protein samples, very small amounts of samples (0.3-0.5 nmol) were used. Since this method is very sensitive to salts, the samples were desalted on a C_4 column (HPLC instrument LC1100, Agilent Technologies, Santa Clara, US-CA) first, then eluted in a combination of 0.1 % TFA (tri-fluoroacetic acid) in H_2O (buffer A) and 0.08 % TFA in acetonitrile (buffer B). A linear gradient from 80 % buffer A and 20 % buffer B to 20 % buffer A and 80 % buffer B was used in the elution. The purpose of adding TFA was to enhance protonation and increase sensitivity. For this only volatile buffers such as ammonium acetate can be used. The MS analysis was carried out on a Finnigan LCQ Advantage MAX mass spectrometer (Thermo Fisher Scientific, Waltham, US-MA) in positive ion mode. The software that we used to analysis the MS spectra was Xcalibur and MagTran. The accuracy of the weight determination of proteins via ESI MS amounted to approximately 7 Da.

2.4.7 Cross-linking analysis coupled with mass spectrometry

10 µM DmMis12a was mixed with BS2G-H6/D6 (Creative Molecules, www.creativemolecules.com) in a molecular weight ratio of 1:1 to a final volume of 50 µL. After incubation at 37 °C for 30 min the reaction was quenched by adding 100 mM ammonium bicarbonate and incubating for 15 min at 37 °C. Cross-linked proteins were digested and the peptides were enriched and analyzed by liquid chromatography coupled to tandem mass spectrometry using a hybrid LTQ-Orbitrap Elite instrument (Thermo Fisher Scientific, Waltham, MA) (Herzog et al., 2012). Cross-links were identified by the

dedicated software xQuest (Walzthoeni et al., 2012). False discovery rates (FDRs) were estimated using xProphet (Walzthoeni et al., 2012) and results were filtered according to the following parameters: FDR = 0.05, min delta score = 0.90, MS1 tolerance window of -4 to 4 ppm, Id-score >22. Cross-links were visualized using the xVis web server (Grimm et al., 2015).

2.4.8 SEC combined with static light scattering

Static light scattering (SLS) can be summarized as a non-invasive method for detecting the molecular mass of a protein sample in solution. By exposure to low intensity laser light (690 nm), the experimental molecular mass of a protein can be determined to an accuracy of better than 5%. The scattered light intensity is measured as a function of angle. And this can be analyzed to produce the molar mass, root mean square radius. In structural studies, the results of SLS experiments are always used to determine the oligomeric state of the solution.

However, this method produces the volume-averaged molecular weight of the sample in the laser beam. And the measurement is very sensitive to baseline errors. So researchers always combine the SEC with the SLS. The light scattering and the concentration are measured for each eluting fraction. For SEC alone, the non-globular proteins may elute at different positions from the predicted, which can be calculated by the calibration curve of the column. If combined with SLS, the molecular mass of all the fractions can be obtained. The results of the SLS are volume averaged. So the purity of the sample is very important to obtain best results. When impure samples are used, there will be increased errors. For the data analysis, the major source of errors is the inaccurate

absorption coefficients. In this study SLS experiments and subsequent data analysis was carried out in collaboration with Dr. Sebastiano Pasqualato.

3 RESULTS and DISCUSSIONS

3.1 Reconstitution of two Mis12 complexes in *D. melanogaster*

As described in the introduction, the human Mis12 complex contains four subunits, whereas in Drosophila one subunit is absent and gene duplication has produced two Nnf1 paralogs. It is important to understand the effect of this gene loss and duplication on the overall structure and function of the Mis12 complex, and to establish the consequences for KMN assembly during mitosis.

In order to determine the organization of the Mis12 complex in *D. melanogaster* (DmMis12), we embarked on an in vitro reconstitution strategy. The DmMis12 complex is composed of four proteins: Mis12 (181 amino acids, 21 kDa), Nnf1a (194 amino acids, 22.4 kDa), Nnf1b (see Figure 9A) (204 amino acids, 23.5 kDa), and Kmn1 (183 amino acids, 21.3 kDa) (Przewloka et al., 2007; Schittenhelm et al., 2007).

To gain the information about the organization of the DmMis12 complex, we expressed all the subunits (see Figure 9A) and their combinations. We first subcloned cDNAs encoding each of the four subunits into pST50Tr-HISDHFR, which encodes a hexahistidine N-terminal tag at the N terminus of the cloning sites (Tan et al., 2005). Mis12, Nnf1a, Nnf1b, and Kmn1 were all insoluble when I expressed them in isolation in *Escherichia coli*. To promote the solubility of these proteins, I tried co-expressing of different combinations of subunits, two at a time, in the pST44 vector (Tan et al., 2005) (see Table 1). $Mis12^{His}$ with Nnf1a or Nnf1b proved to be soluble complexes, while binary combinations containing Kmn1 were insoluble (see Table 1 and Figure 9).

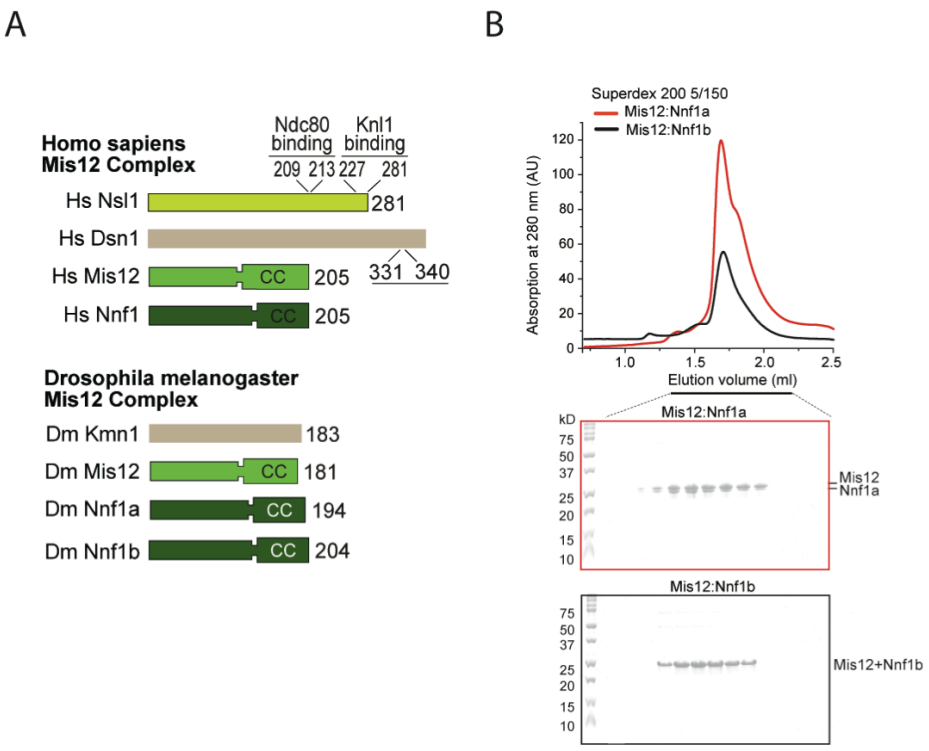

Figure 9 Different sub-complexes of DmMis12 complex

(**A**) Schematic presentation of the constitutive subunits of the Mis12 complex in humans and in *Drosophila*. Segments identified for their ability to interact with Knl1 or Ndc80 complex subunits are indicated in the text; (**B**) Size-exclusion chromatography of binary complexes of Mis12 and Nnf1a or Nnf1b and SDS-PAGE analysis of showed fractions.

These results suggest that Mis12 and Nnf1 interact extensively within the *Drosophila* Mis12 complex, as also observed with the Mtw1/Mis12 complex of *Saccharomyces cerevisiae* (Maskell et al., 2010b). SEC elution profiles of $Mis12^{His}$: Nnf1a and $Mis12^{His}$: Nnf1b sub-complexes are shown Figure 9B. SDS-PAGE analysis revealed that the $Mi12^{His}$ and Nnf1a ran close together on the gel, while $Mis12^{His}$ and Nnf1b were overlaid on the gel (see Figure 9B). This is to be expected because, as previously stated, the molecular weights of these proteins are very similar: Mis12, 21 kDa; Nnf1a, 22.4 kDa; Nnf1b, 23.5 kDa.

Combinations containing three ($Mis12^{His}$: Nnf1a: Kmn1, $Mis12^{His}$: Nnf1b: Kmn1) or four ($Mis12^{His}$: Nnf1a: Nnf1b: Kmn1) subunits were also obtained after appropriate

sub-cloning into the pST44 system. Solubility of Kmn1 was only observed when it was co-expressed in combination with Mis12 and either Nnf1a or Nnf1b. All of these expression test results are summarized in Table 1.

Since Kmn1 can only be solubilized with Nnf1a and/or Nnf1b, I hypothesized that the four subunits of the DmMis12 complex may have a similar organization with the human Mis12 complex, in which the four subunits form a tight complex (Petrovic et al., 2010b).To test this hypothesis, we purified the $Mis12^{His}$:Nnf1a:Nnf1b:Kmn1 complex (see Figure 10A). The profile was mono-disperse, and the purity from the SDS-PAGE was more than 95%. Moreover, the samples appeared homogeneous. In order to investigate the stoichiometry of the Mis12 complex, we then tested the sample with static light scattering (SLS) combined with SEC (see Figure 10B).

The green profile in Figure 10B is the Right Angle Light scattering signal and it can be seen that the peak visible at UV was almost symmetrical. The regions of slices 4 and 5 were reliable for data analysis. This part is homogeneous, with a molecular weight of 60 to 62 kDa. As we showed before, the approximate mass of each individual subunit is 20 KDa. These SLS results are therefore incompatible with the existence of a tetramer, and are a better fit with a three-protein complex, indicating that a stable three subunits complex may exist when $Mis12^{His}$, Nnf1a, Nnf1b, and Kmn1 are co-expressed.

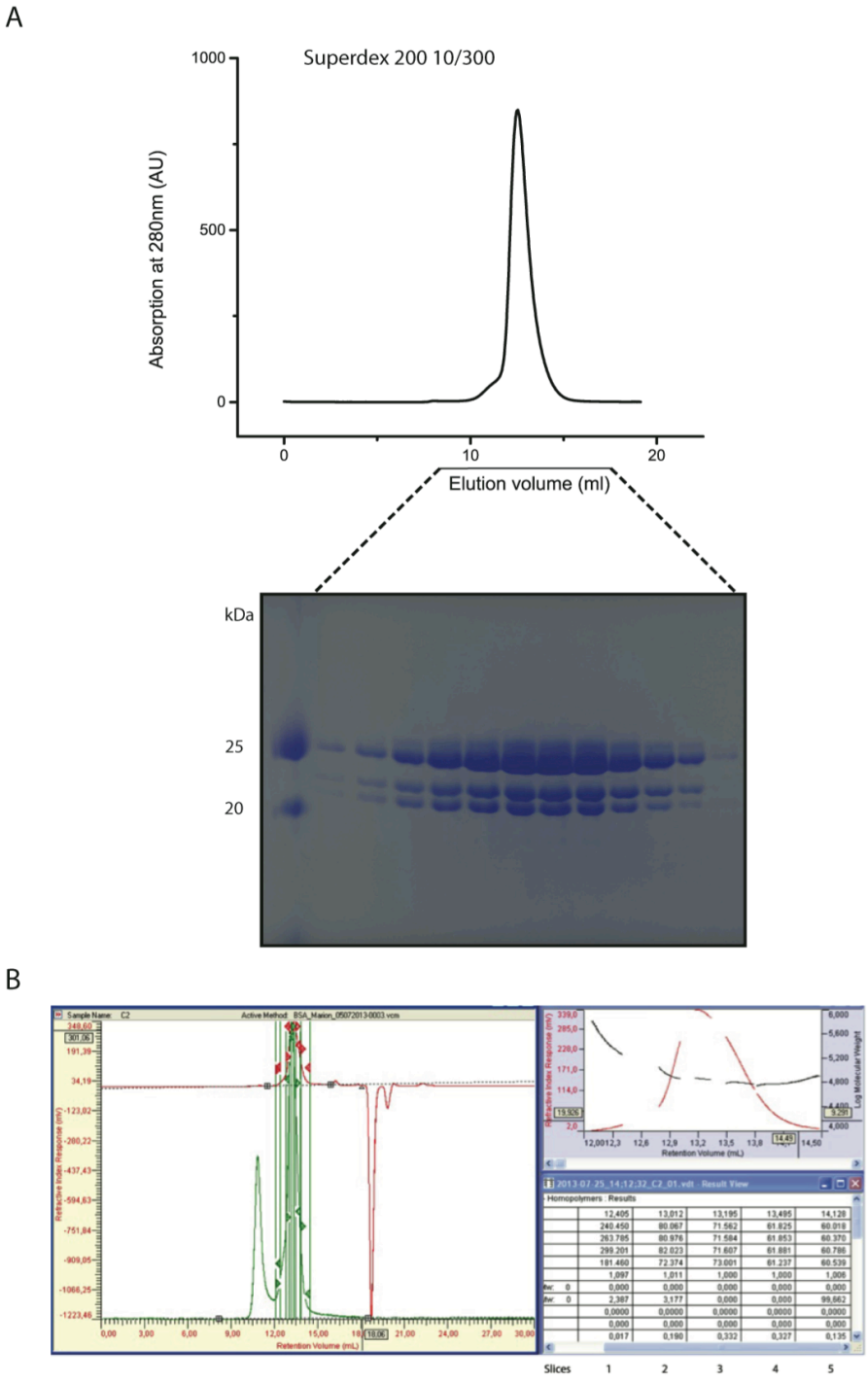

Figure 10 Purification and characterization of Mis12: Nnf1a: Nnf1b: Kmn1 complex

(**A**) Size-exclusion chromatography of complexes obtained by co-expression of Mis12, Nnf1a, Nnf1b, and Kmn1, and SDS-PAGE analysis.

(**B**) SEC combined with static light scattering analysis of complexes obtained by co-expression of Mis12, Nnf1a, Nnf1b, and Kmn1, and SDS-PAGE analysis.

Table 1 Summary of the expression tests of Mis12C in Drosophila

"Soluble" or "Insoluble" indicates that the protein could/could not be identified in the bacterial cell lysate.

	Monomers				**Dimers**		**Trimers**	**Tetramer**
Construct	Mis12	Nnf1a	Nnf1b	Kmn1	Mis12: Kmn1 Kmn1: Nnf1a Kmn1: Nnf1b	Mis12: Nnf1a Mis12: Nnf1b	Mis12: Nnf1a: Kmn1 Mis12: Nnf1b: Kmn1	Mis12: Nnf1a: Nnf1b: Kmn1
Solubility	Insoluble	Insoluble	Insoluble	Insoluble	Insoluble	Soluble	Soluble	Soluble

Since the SLS data indicated a three subunits complex, it seems possible that Nnf1a and Nnf1b are incorporated in different complexes. To address this question, we co-expressed Mis12, Nnf1a, Nnf1b, and Kmn1, with each subunit fused to a different tag: Mis12 with a Flag tag, Kmn1 with an Hpc tag, Nnf1a with a Strep tag, and Nnf1b with a His tag. Cell lysates were pelleted and the cleared lysate was incubated with resins targeting the affinity tags of Nnf1a (Strep tag) and Nnf1b (His tag) in consecutive steps (see Figure 11). To remove the unspecific binding before eluting the protein complexes from the columns, both columns were washed extensively. After elution, the bound fraction from each step was analyzed using Western blotting (WB) (see Figure 11). We suspected that the elution from the StrepTrap™ column would release a complex containing Nnf1a, and that the elution from the His Trap™ column would release a complex containing Nnf1b. The WB revealed that Mis12 and Kmn1 exist in both Nnf1a and Nnf1b complexes, while in the Nnf1a complex no Nnf1b was present. Conversely, in the Nnf1b complex, no Nnf1a was present. Thus, Nnf1a and Nnf1b form distinct three-subunit complexes with Kmn1 and Mis12, which I define separately as the DmMis12a and DmMis12b complexes.

We reconstituted the DmMis12a and the DmMis12b complexes and purified them after bacterial co-expression (see Methods section). In both complexes, only Mis12 was tagged with 6 histidines. The size-exclusion chromatography (SEC) purification step, which separates proteins based on their shape and molecular mass, demonstrated that both complexes are monodisperse and overlaying the profiles indicated that the two complexes have a similar shape and overall mass (see Figure 12A).

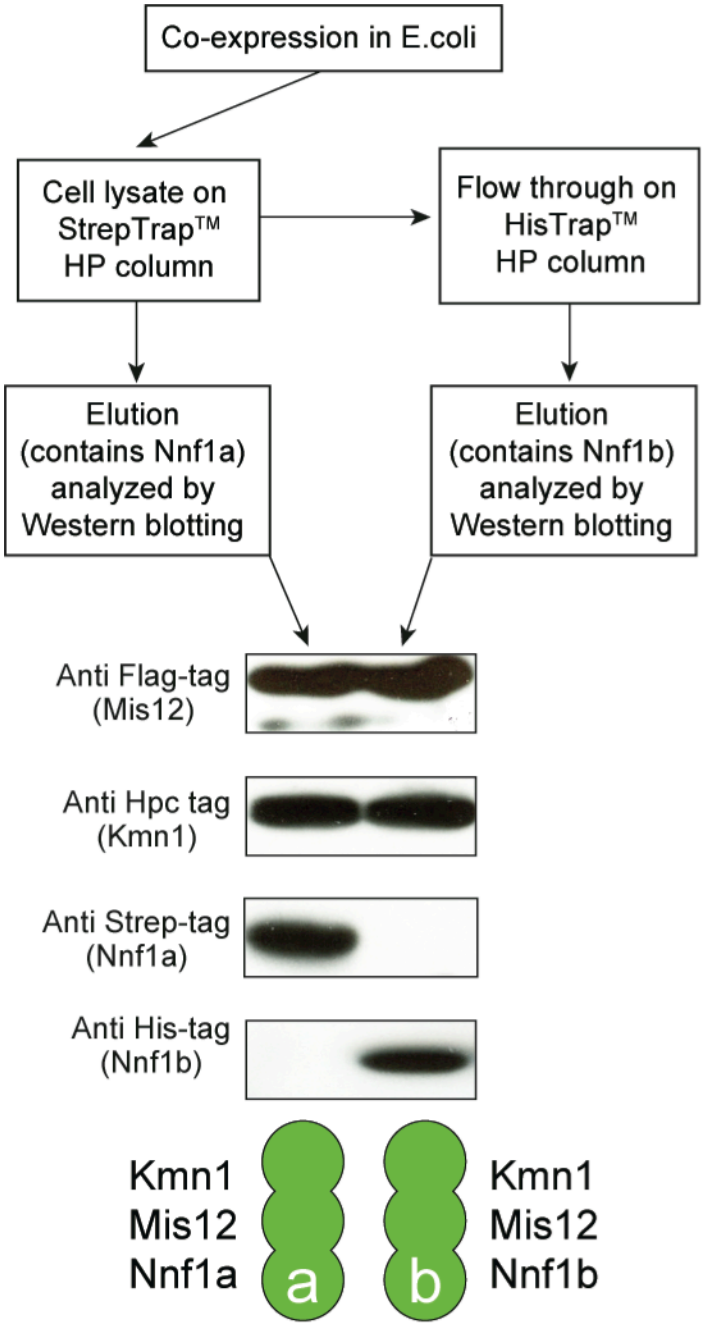

Figure 11 Two Mis12 complexes in Drosophila Melanogaster

The strategy used to determine whether Nnf1a and Nnf1b are part of the same or different complexes is displayed here. In order to eliminate false positive results, the StrepTrap™ column and HisTrap™ column were washed extensively before elution.

The SEC profile of the complexes generated by co-expression of four genes exhibited a strong overlap with the DmMis12a and DmMis12b complexes (see Figure

12A), which can be interpreted as the sum of the two peaks corresponding to two different Mis12 complexes, containing either Nnf1a or Nnf1b.

As already discussed, previous work indicates that the Mis12 complex may be subdivided into two sub-complexes: one of Mis12-Nnf1 and one of Dsn1-Nsl1 in humans and yeast (Maskell et al., 2010a; Petrovic et al., 2010b). Using a reconstitution experiment, we determined that Mis12 and Nnf1a/Nnf1b also form a tight and constitutive dimer. Furthermore, two Mis12 complexes (DmMis12a and DmMis12b) exist, although no direct interaction between these two independent complexes takes place. In the next section I will describe detailed functional analysis of the two complexes that will bring to light how different subunits interact with each other.

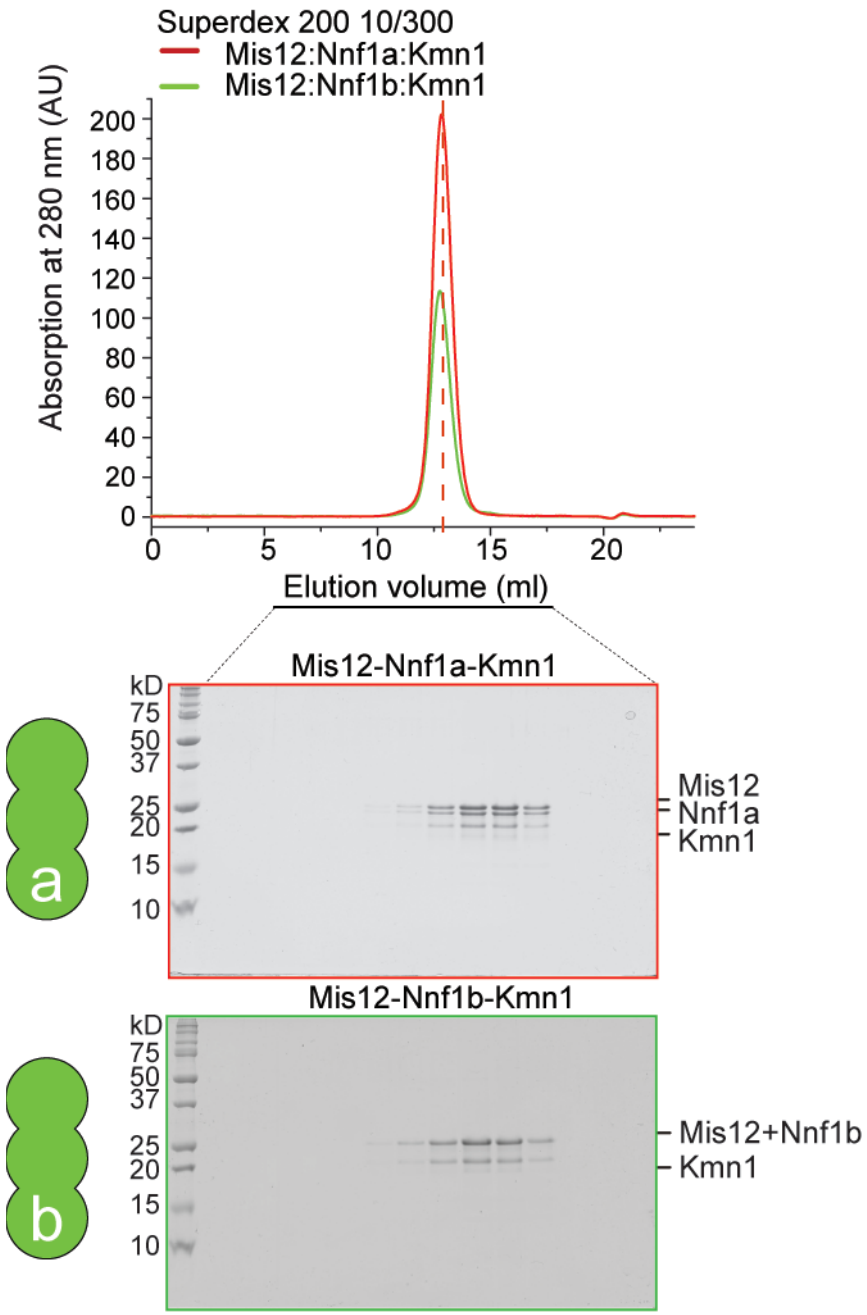

Figure 12 Size-exclusion chromatography (SEC) experiment on the DmMis12a and DmMis12b complex

Both complexes elute in a single peak and appear mono-disperse. The vertical dashed bar is a reference indicating the elution volume of the complexes generated by co-expression of four gene products.

3.2 Characterization of the two Mis12 complexes

Analytical ultracentrifugation (AUC) sedimentation velocity experiments were used to detect the molecular weight of the complexes (the results are summarized in Table 2). These experiments indicated that the stoichiometry of both the DmMis12a and DmMis12b complexes is 1:1:1, which agrees with the results of the SLS experiments in Figure 10. In general, for a globular protein, the frictional ratio is around 1.2 (Garbett et al., 2010). In our case, DmMis12a and DmMis12b complexes both have frictional ratios of approximately 1.7, suggesting that both complexes are elongated.

In order to visualize the overall structure of the DmMis12 complexes, negative-stain electron microscopy (EM) experiments were carried out on the DmMis12a complex (see Figure 14A). The results showed that the majority of single particles appeared elongated, that one end is thicker in comparison to the other, and that the overall length of the complex is ~20 nm. This overall structure of the DmMis12a complex is comparable to previous EM studies on human and budding yeast Mis12 complexes (see Figure 14B-C) (Hornung et al., 2011b; Maskell et al., 2010b; Petrovic et al., 2014b; Petrovic et al., 2010a). Consequently, the loss of one subunit does not dramatically alter the elongated appearance of the DmMis12 complex.

SEC experiments lead to a high purity and monodispersity of the DmMis12 complexes (see Figure 12A). However, under negative stain EM, we observed more structural heterogeneity (see Figure 14A), which prevented us from obtaining stable class averages.

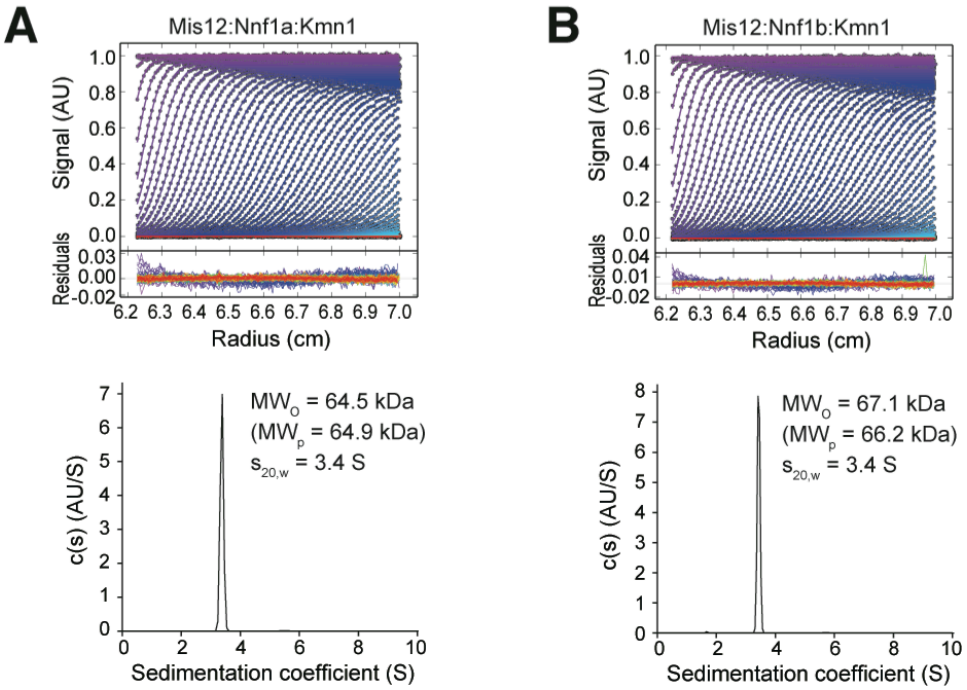

Figure 13 AUC experiments of DmMis12a and DmMis12b complexes

Sedimentation velocity absorbance profiles of the DmMis12a (**A**) and DmMis12b (**B**) complexes, with residuals of the fit showing the deviation of the c(*S*) model from the observed signals; the best-fit continuous-size c(*S*) distribution of the DmMis12a and DmMis12b complexes is shown in the bottom section of the panel.

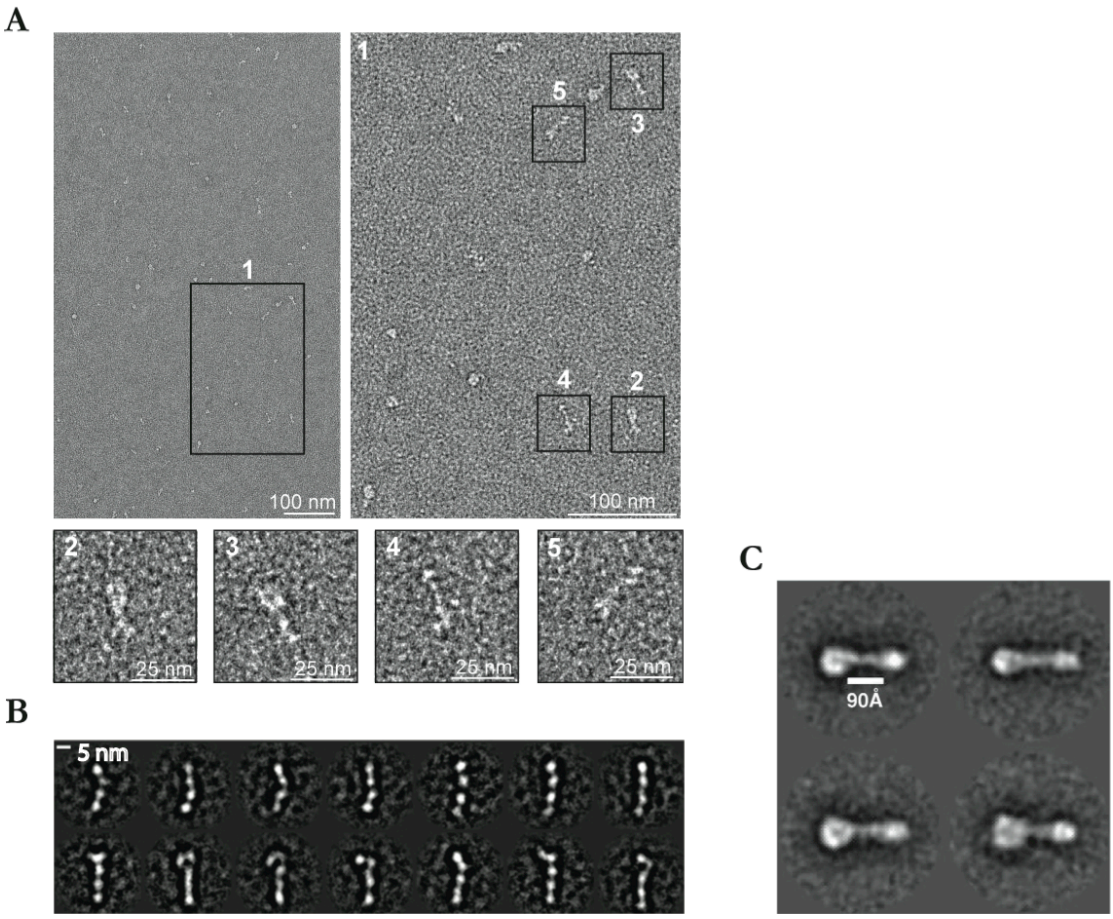

Figure 14 Representative negative stain EM images of the Mis12 complex

(**A**) The DmMis12a complex is very elongated under the EM. In comparison to the Mis12 complex in humans, it is more structurally heterogeneous; (**B**) the class averages represent the characteristic views of the Mis12 complex in humans; (**C**) the class averages showing the architecture and dimensions of the Mis12 complex in budding yeast; The scale bars are indicated in the figures.

We set out to test which parts of the Mis12 complex are susceptible to structural heterogeneity in order to generate samples that may be amenable for further structural studies. SDS-PAGE data in Figure 12A showed that there were clear degradation bands in both the DmMis12a and DmMis12b complexes. These indicated the potential presence of flexible or unstable parts in both complexes. To obtain stable core complexes for further structural study, small scale limited proteolysis experiments were carried out for both complexes (see Figure 15).

Four proteases were utilized in this experiment: Papain, Trypsin, Subtilisin and Chymotrypsin. For both the DmMis12a and the DmMis12b complex, we could see that Trypsin and Chymotrypsin generated similar stable bands (Red arrows in Figure 15). Mis12, Nnf1a/Nnf1b were relatively stable, whereas the Kmn1 appeared to disappear in both complexes. This suggests that the degradation bands stem from Kmn1 (see Figure 15).

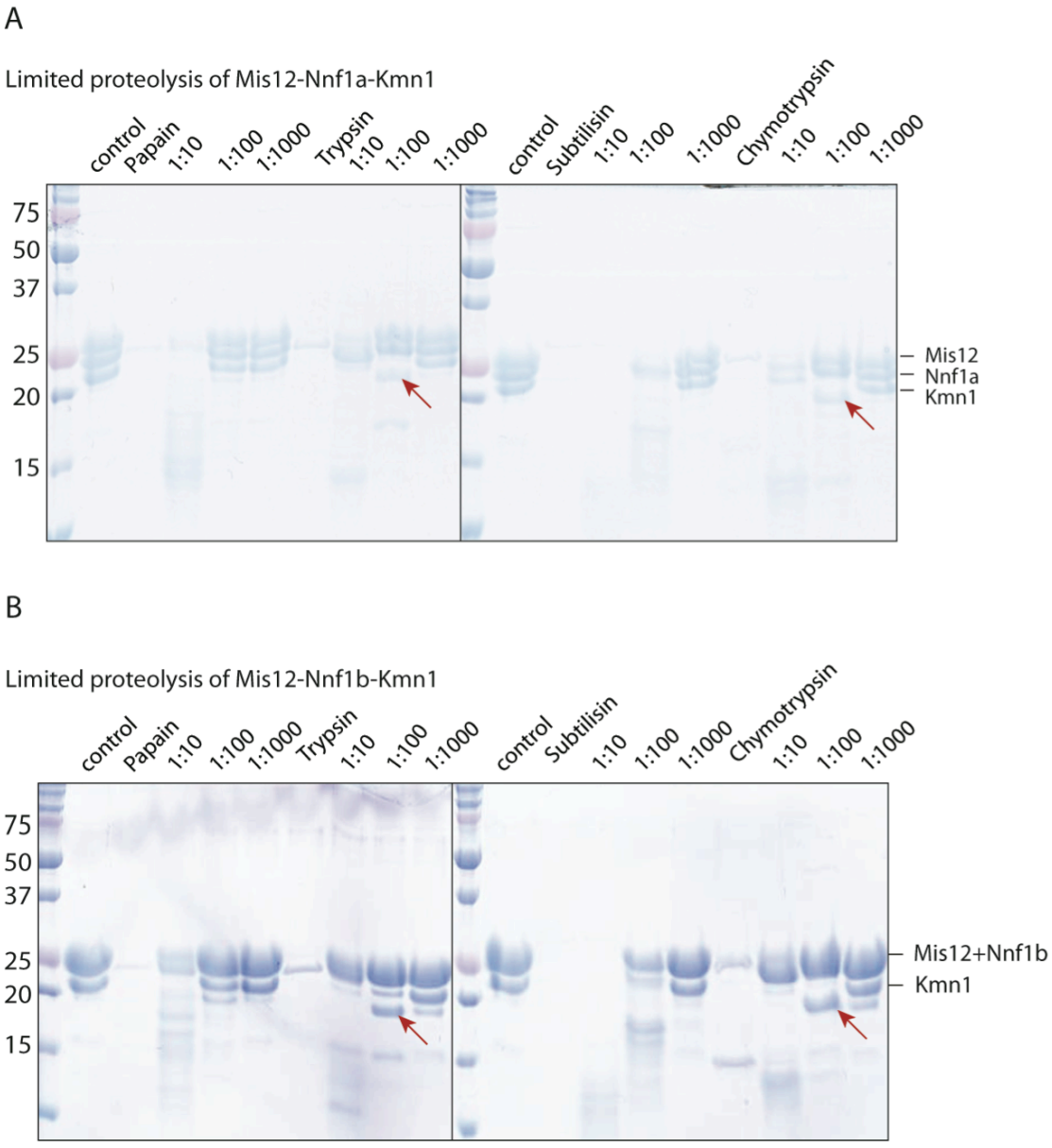

Figure 15 limited proteolysis of DmMis12a and DmMis12b complexes

Methods are detailed in the method section. Compared with the control (the native complexes), in both DmMis12a (**A**) and DmMis12b (**B**) complexes, Trypsin and Chymotrypsin can eliminate Kmn1 and generate similar products, as indicated by the red arrows. The other subunits, Mi12, Nnf1a/Nnf1b proved to be relatively stable.

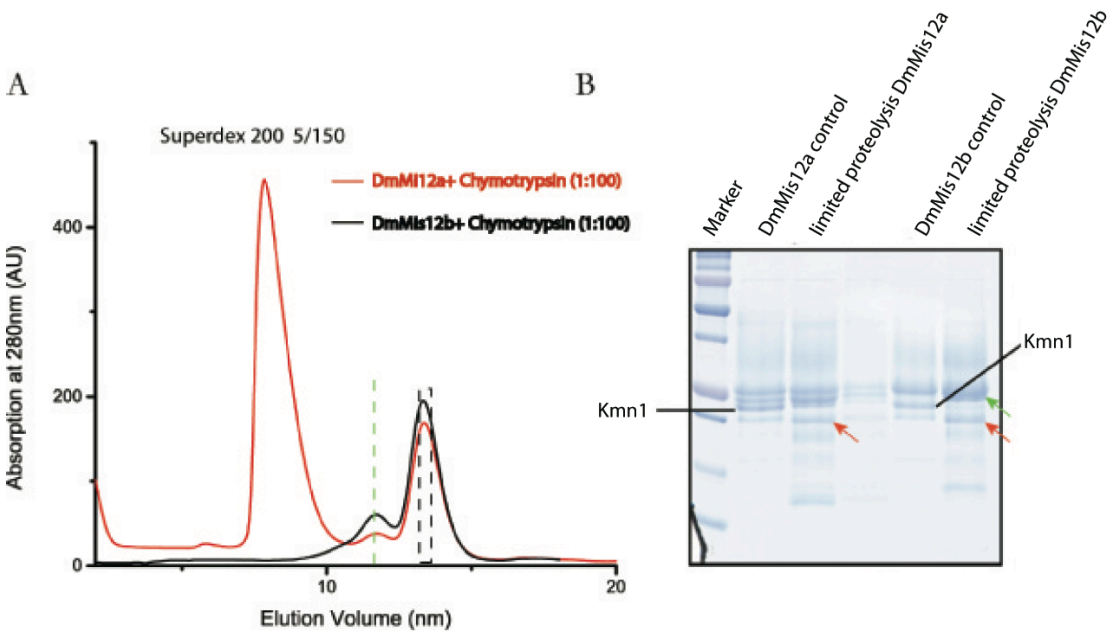

Figure 16 Scaled up limited proteolysis of DmMis12a and DmMis12b complexes

(**A**) SEC of the DmMi12a/DmMis12b with Chymotrypsin. The green dashed line shows the elution position of native DmMis12a and DmMis12b, which indicates there are still un-digested DmMis12C. The dash black box shows the fraction that we used for SDS-PAGE and Electrospray ionization mass spectrometry (ESI-MS). (**B**) The red arrow illustrates the stable fragments (similar to Figure 15). This indicates that in DmMis12b, Mis12 and/or Nnf1b had been partially digested.

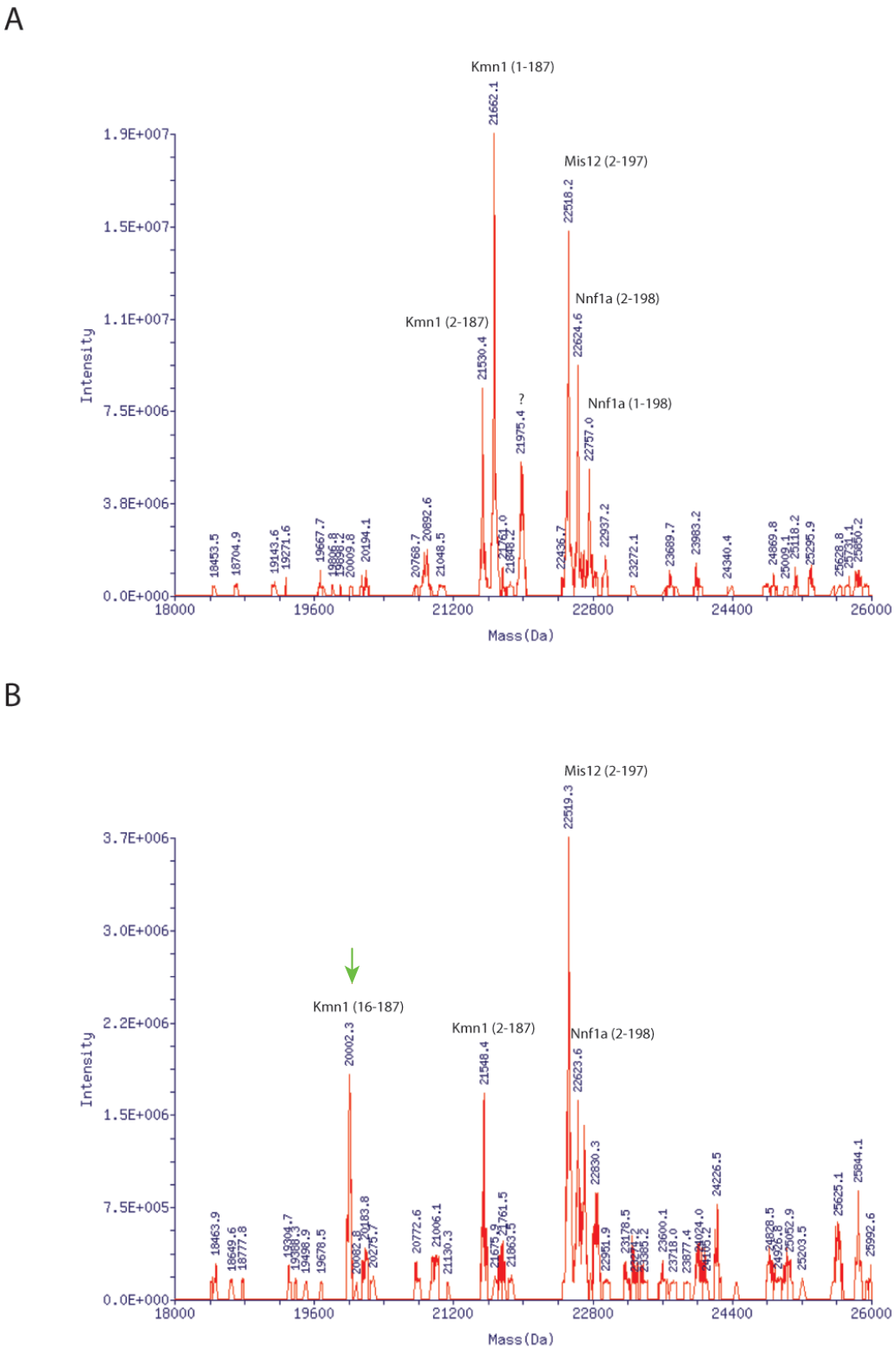

Figure 17 ESI mass spectrometry of DmMis12a complex

(**A**) The control sample, which contains the native Dm Mis12a complex. (**B**) Compared to the Native DmMis12a complex, the processed (by Chymotrypsin) sample exhibited a new peak: Kmn1 (16-187) (indicated by the green arrow).

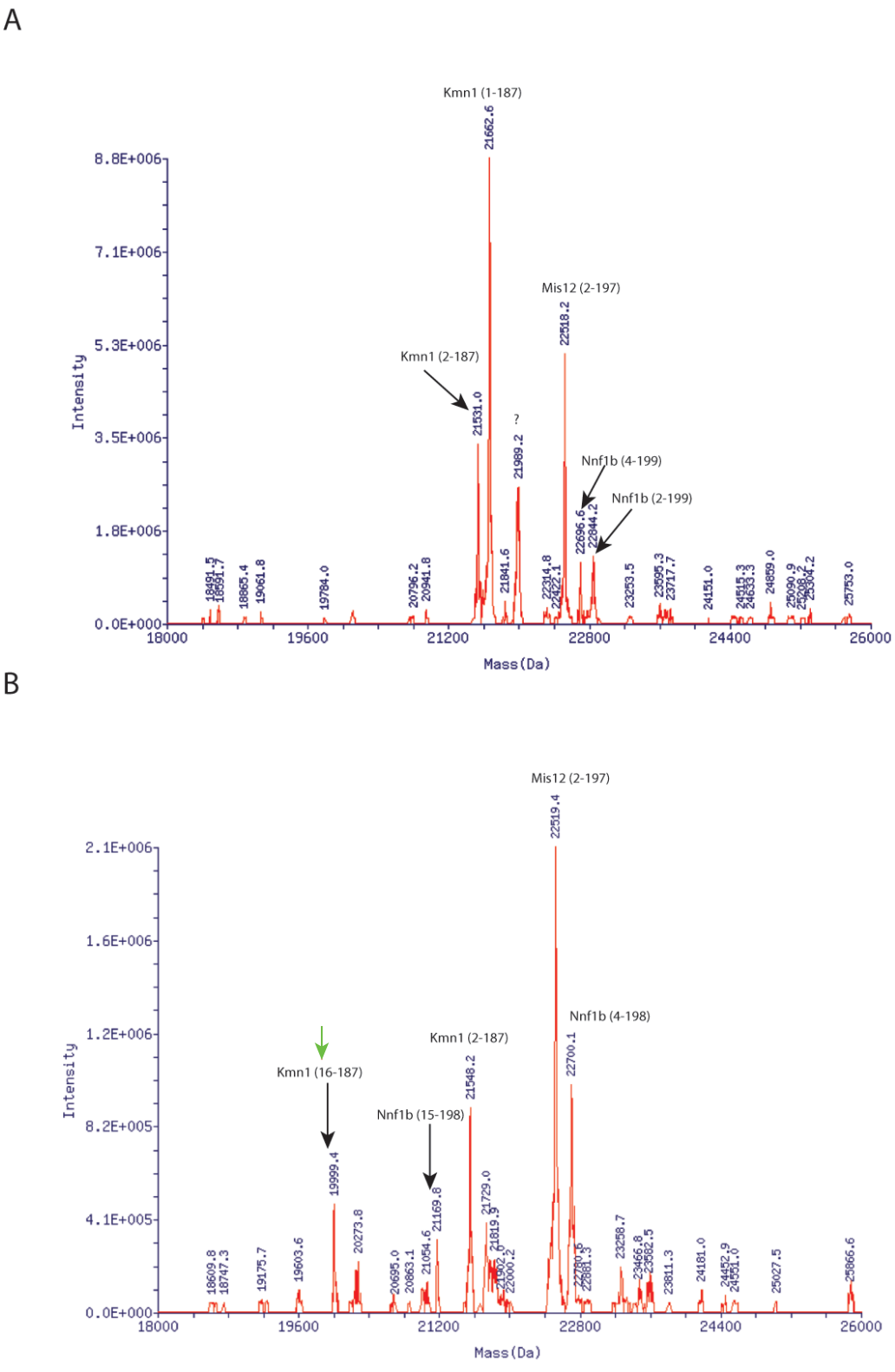

Figure 18 ESI mass spectrometry of DmMis12b complex

(**A**) The control sample, which contains the native DmMis12b complex. (**B**) Compared to the Native DmMis12b complex, the processed (by Chymotrypsin) sample exhibited a new peak: Kmn1 (16-187) (indicated by the green arrow) and Nnf1b (4-198).

Table 2 Summary of sedimentation velocity experiments. All predicted molecular masses assumed each subunit was present in a single copy

Complexes	Frictional Radio	Observed molecular mass (kDa)	Predicted molecular mass (kDa)	S (20,w)
Mis12: Nnf1a: Kmn1	1.7	64.5	64.9	3.4
Mis12: Nnf1b: Kmn1	1.7	67.1	66.2	3.4
Mis12: Nnf1a: Kmn1: Cenp-C$^{1-105}$	1.9	76.4	76.8	3.4
Mis12: Nnf1b: Kmn1: Cenp-C$^{1-105}$	1.9	75.6	78.0	3.4
Mis12: Nnf1a: Kmn1: Spc105R$^{1707-1960}$	1.7	93.3	95.2	4.3
Mis12: Nnf1b: Kmn1: Spc105R$^{1707-1960}$	1.6	95.7	96.4	4.8
Mis12: Nnf1b: Kmn1: Spc105R$^{1707-1960}$: Cenp-C$^{1-105}$	1.8	106.5	109.3	4.4

Table 3 Summary of deletion mutants of the loop regions in DmMis12a and DmMis12b complexes

Complex	Proteins	Target Loops	Deletions	Complex stability
Mis12-Nnf1a-Kmn1	Mi12	Loop1	Mis12$^{\Delta 85-95}$	Unstable
			Mis12$^{\Delta 87-93}$	Unstable
			Mis12$^{\Delta 89-92}$	Unstable
	Nnf1a	Loop2A	Nnf1a$^{\Delta 101-109}$	Unstable
			Nnf1a$^{\Delta 103-107}$	Unstable
	Kmn1	Loop3	Kmn1$^{\Delta 29-41}$	Stable
			Kmn1$^{\Delta 32-38}$	Stable
		Loop4	Kmn1$^{\Delta 58-70}$	Stable
			Kmn1$^{\Delta 61-66}$	Stable
Mis12-Nnf1b-Kmn1	Mis12	Loop1	Mis12$^{\Delta 85-95}$	Unstable
			Mis12$^{\Delta 87-93}$	Unstable
			Mis12$^{\Delta 89-92}$	Unstable
	Nnf1b	Loop2B	Nnf1b$^{\Delta 107-113}$	Unstable
			Nnf1b$^{\Delta 109-112}$	Unstable
	Kmn1	Loop3	Kmn1$^{\Delta 29-41}$	Stable
			Kmn1$^{\Delta 32-38}$	Stable
		Loop4	Kmn1$^{\Delta 58-70}$	Stable
			Kmn1$^{\Delta 61-66}$	Stable

We attempted to scale up these experiments using only Chymotrypsin (1:100 dilution), and used SEC to separate the protease from the DmMis12a and DmMis12b complexes following incubation with the diluted protease (see Figure 16). From the SEC profiles, we can see the digested complexes shift to the right (see Figure 16A). However, despite the large changes in the elution volume, Mis12 complex remains intact as seen from the presence of three bands on the SDS-PAGE. (see Figure 16B).

In order to identify the stable fragments after limited proteolysis, ESI mass spectrometry was used to test the samples from SEC (see Figure 17 and Figure 18). In both the DmMis12a and the DmMis12b complex, the ESI MS confirmed that 16 amino acids had been removed form the N terminus of Kmn1 (DmMis12C$^{kmn1\Delta N16}$). In the DmMis12b complex, in addition to Kmn1, Nnf1b had also lost 4 amino acids from its N terminus. In both cases, similar peaks to the control were evident. For example, in Figure 17, the peaks of the partially digested DmMis12a and Nnf1a were the same as those of the control sample, implying undigested complexes remained. From the ESI MS results, we concluded that part of the N terminal of Kmn1 is flexible, and can thus be digested by Chymotrypsin without affecting the overall stability of both the DmMis12a and DmMis12b complexes.

Despite very extensive attempts, we were unable to obtain crystals of the full length DmMis12 complexes or of the DmMi12C$^{kmn1\Delta N16}$ complex identified through limited proteolysis (data not shown).In order to acquire additional information about all potentially flexible parts of both complexes, we examined secondary structure predictions of all the four subunits, as shown in Figure 19. Our focus was on the regions of more than 10 amino acids predicted to be unstructured (indicated by black arrows in Figure 19): loop 1 in Mis12; loop 2A in Nnf1a; loop 2B in Nnf1b; loop 3 and 4 in Kmn1. Based on this analysis, we made a new series of deletion mutants, as summarized in Table 3. The deletion of Loop1 in Mis12, Loop2A in Nnf1a and Loop2B in Nnf1b lad

to destruction of the complexes, implying that these segments are all essential for complex stability. Deletions as short as 3 or 4 amino acids resulted in the total destabilization of the complex, as argued based on the fact that these complexes cannot be expressed in, or purified from, *E.coli* (see Table 3). However, Loop 3 and 4 at the N terminal of Kmn1 appeared to be quite flexible as deleting them did not affect the stability of either the DmMis12a at the DmMis12b complex. These deletions provided confirmation that the N terminal of Kmn1 contains a flexible region or an independent domain.

In order to perform crystallization, we optimized different constructs of the N terminal of Kmn1 and also of the DmMis12a complex with different parts of Kmn1. We purified the N terminal sections of a Kmn1 (1-68) fragment (which we named the Head region) and the DmMis12a complex without Head region (termed the Delta DmMis12a complex), as shown in Figure 20. This revealed that the Head region in the DmMis12a complex is relatively independent, and would not affect the overall stability of the DmMis12a complex. Unfortunately, neither the Head region nor the Delta DmMis12a complex yields crystals. However, we used Circular dichroism (CD) to test the secondary structure of this Head region (See Figure 20C). Secondary structure analysis software such as CDNN indicates that about 70% of these regions are unstructured (http://bioinformatik.biochemtech.uni-halle.de/cdnn).

Table 4 Summary of expression experiments with different deletion mutants of the subunits of the DmMis12a complex

Complexes	Deletions	Complex Stability	Subunit Stability
Mis12: Nnf1a	Mis12$^{\Delta 100-181}$	Low	Mis12$^{\Delta 100-181}$ soluble Nnf1a insoluble
	Nnf1a$^{\Delta 181-194}$	Normal	----
	Nnf1a$^{\Delta 124-194}$	Low	Nnf1a$^{\Delta 124-194}$ soluble Mis12 insoluble
Mis12: Nnf1a: Kmn1	Kmn1$^{\Delta 130-183}$	Low	Mis12: Nnf1a soluble Kmn1$^{\Delta 130-183}$ insoluble

Table 5 Summary of expression experiments with different deletion mutants of the subunits of the DmMis12b complex

Complexes	Deletions	Complex Stability	Subunit Stability
Mis12: Nnf1b	Mis12$^{\Delta 100-181}$	Low	Mis12$^{\Delta 100-181}$ soluble Nnf1b insoluble
	Nnf1b$^{\Delta 129-208}$	Low	Nnf1b$^{\Delta 129-208}$ soluble Mis12 insoluble
Mis12: Nnf1b: Kmn1	Kmn1$^{\Delta 130-183}$	Low	Mis12: Nnf1b soluble Kmn1$^{\Delta 130-183}$ insoluble

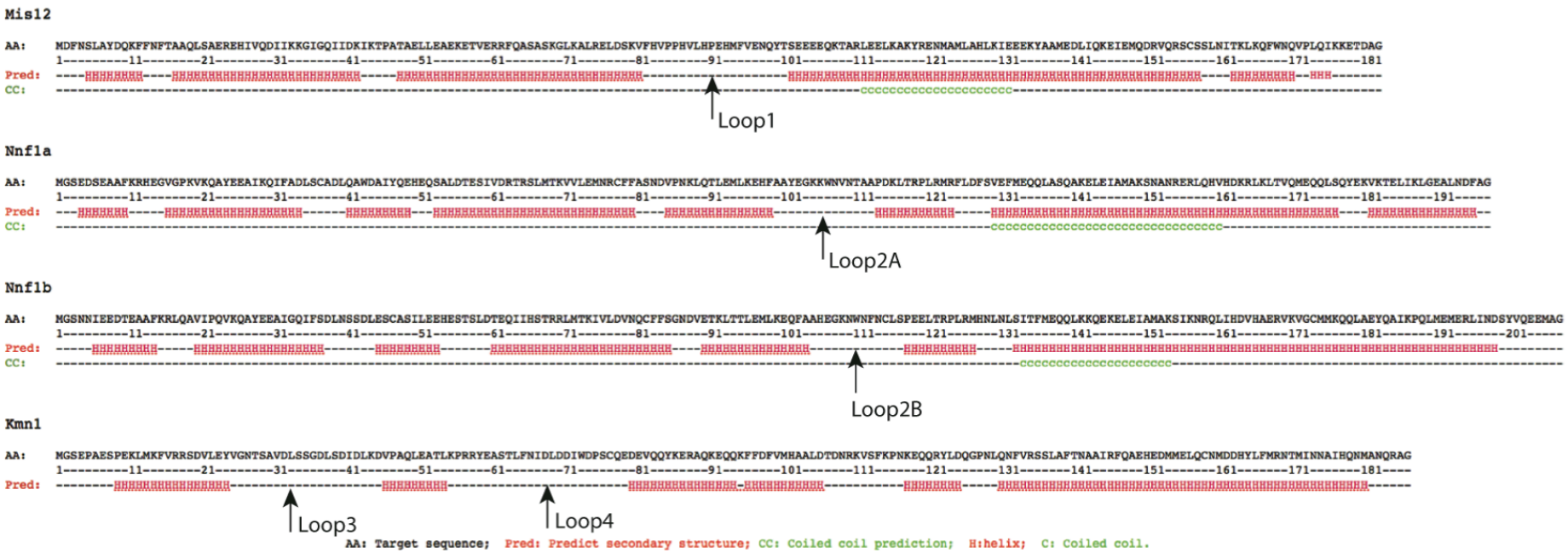

Figure 19 Second structural predictions of DmMis12a and DmMis12b complexes

The red H indicates the predicted helical folding. Between the red Hs, there are small dashes indicating the predicted unstructured regions (we call them loop regions). The green C represents the predicted coiled coil region. We can see that there are coiled coil at the C terminal of Mis12, Nnf1a and Nnf1b. Jpred 4 were used for the second structural prediction (Drozdetskiy et al., 2015).

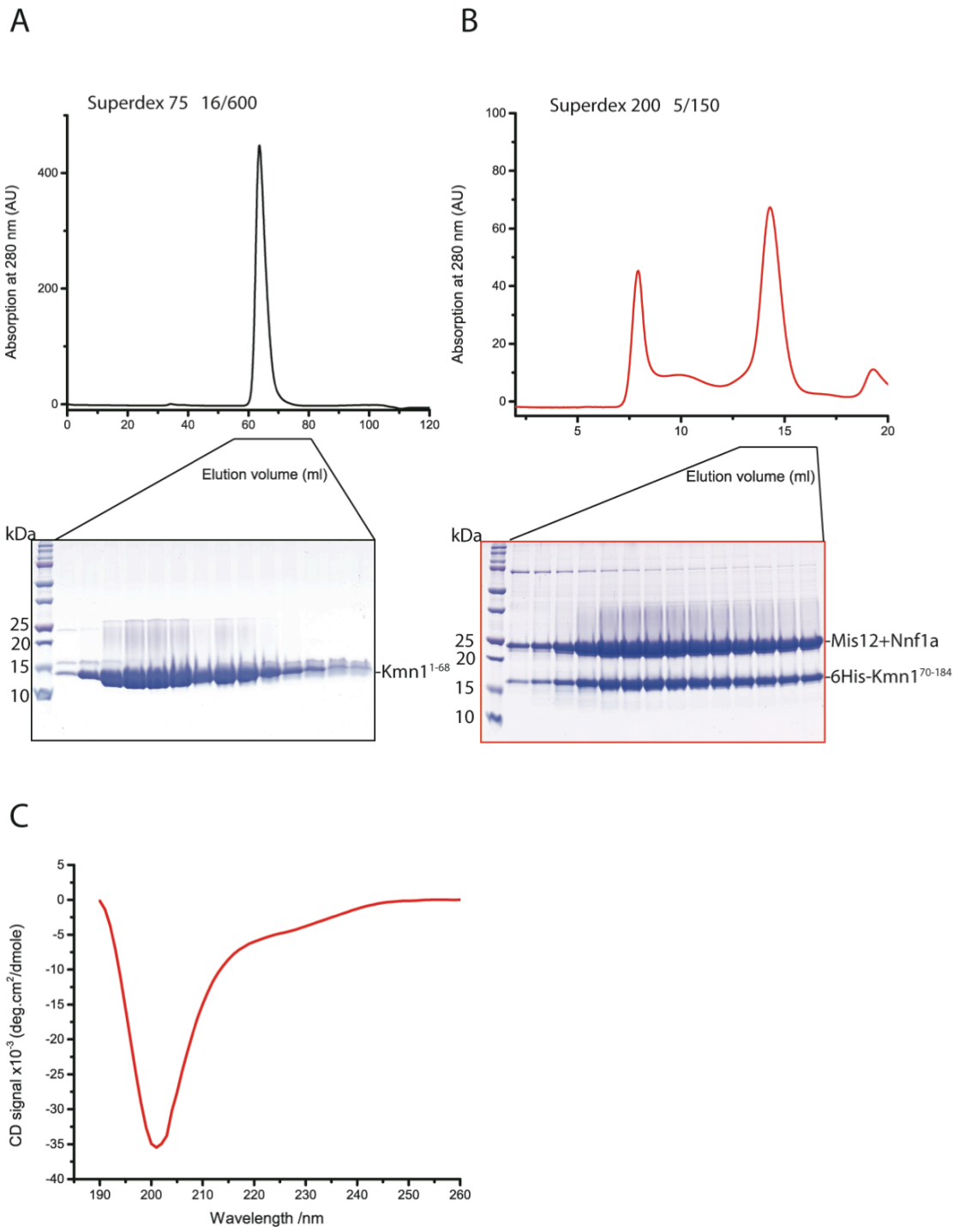

Figure 20 Purification of N terminal Kmn1 and Delta DmMis12a complex

Size-exclusion chromatography (SEC) experiments on Kmn1$^{1-68}$ (**A**) and Delta DmMis12a complex (**B**) show the fragment and the protein complex elute in a single peak and appear monodisperse; (**C**) Circular dichroism (CD) spectra results of Kmn1$^{1-68}$.

Extensive deletion experiments for crystallization are summarized in Table 4. These revealed that residues 130 to 183 in the C-terminal region were required for a stable interaction of Kmn1 with the rest of the DmMis12a complex, as their deletion (Kmn1Δ130-183) generated an unstable mutant that failed to be incorporated in a complex with Nnf1a and Mis12. Large C-terminal deletions of Mis12 and Nnf1a also strongly reduced the stability of the binary Mis12:Nnf1a complex. Since the second structural prediction suggests that the C terminal of both Mis12 and Nnf1a have long coiled coil regions, we assume that the coiled coil plays an important role in stabilizing the binary Mis12: Nnf1a or ternary Mis12: Nnf1a: Kmn1 complexes. Results were similar for the DmMis12b complex (see Table 5).

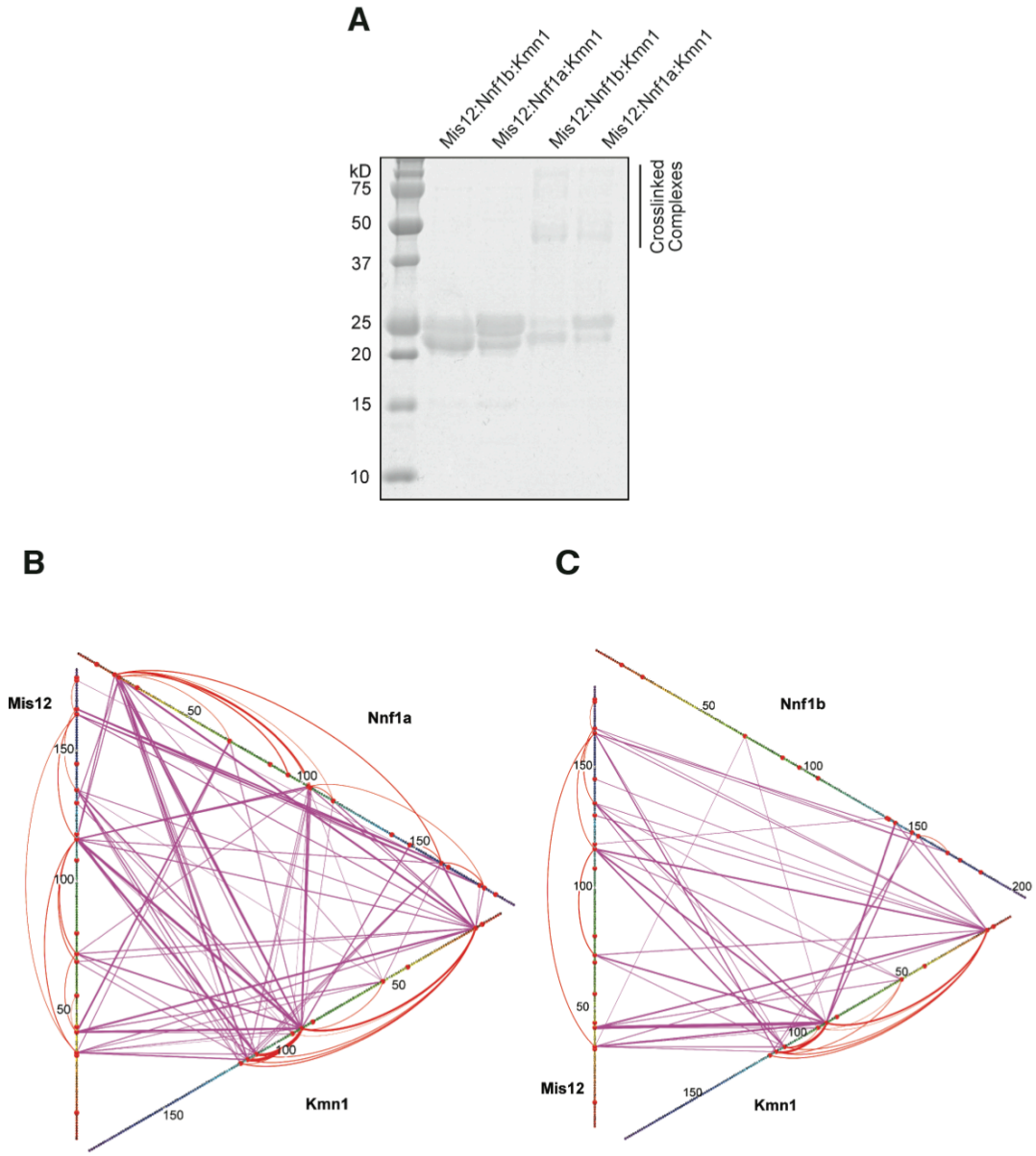

Figure 21 Cross-linking mass spectrometry (XL-MS) analysis of the DmMis12a and DmMis12b complexes

(**A**) Cross-linking of the indicated samples with BS2G (bis[sulfo-succinimidyl]glutarate) leads to the accumulation of high molecular weight species that were subsequently subjected to mass spectrometry analysis. (**B**) XL-MS analysis results of DmMis12a complex. (**C**) XL-MS analysis results of DmMis12b complex. Blue and red lines indicate inter- and intra-molecular cross-links, respectively.

Chemical cross-linking with the bi-functional reagent BS2G (bis[sulfo-succinimidyl]glutarate)(see Figure 21A), combined with protease digestion and mass

spectrometry (XL-MS) (Herzog et al., 2012), was used to gain additional insights into the organization of the DmMis12a and DmMis12b complexes. The results suggest a very extensive interaction between the Mis12 and Nnf1a or Nnf1b subunits, which extends along the full length of their sequences (see Figure 21B and Figure 21C). Combined with the elongated appearance showed in Figure 14, it seems reasonable to suppose that the overall structure of all the subunits may not be globular, but elongated in the complexes.

In summary, using Western Blotting, our biochemical reconstitution data first demonstrated that Nnf1a and Nnf1b were in different complexes in vitro (see Figure 11). We identified two distinct *Drosophila* Mis12 complexes, containing either the Nnf1a or the Nnf1b subunit, which we called DmMis12a and DmMis12b. SLS and AUC studies indicated the stoichiometry of both complexes was 1:1:1, and this result further implied that the three-subunit Mis12C could still function like the four-subunit counterparts in humans and yeast. The negative stain electron microscopy study indicated that the DmMis12 complexes are elongated rod like proteins with one end thicker than the other (see Figure 14A). This feature is similar to the Mis12 complex in humans (see Figure 14B-C) (Petrovic et al., 2010b) and yeast (Hornung et al., 2011a; Maskell et al., 2010a). Even the length of the DmMis12C rod is similar: In *Drosophila*, the length is approximately 20 nm (see Figure 14A) as the counterpart in humans (Petrovic et al., 2016; Petrovic et al., 2010b). This implies that even though the Mis12 complex in *Drosophila* has lost one subunit and the existing subunits contains sequences that markedly diverge from those in humans and yeast, the overall structure is generally conserved during evolution.

3.3 DmMis12-C interacts directly with CENP-C

CENP-C, the only conserved subunit of the CCAN complex in both *Drosophila* and humans, interacts directly with the CENP-A nucleosome in the centromere chromatin with the C terminus (as summarized in the Introduction). Figure 6 compares the overall organization of CENP-C in *Drosophila melanogaster* with its human homolog. Remarkable differences are evident. For example, DmCENP-C contains ~500 residues more than Hs CENP-C (Heeger et al., 2005b). The Arginine-rich (R-rich) domain and the Drosophilids CENP-C homology (DH) domain are unique domains (Heeger et al., 2005b), whose functions are still unknown. At the equivalent position of HsCENP-C, there is a region that has recently been reported to be responsible for binding to CCAN subunits, such as CENP-HIKLMN, a CCAN complex containing CENP-LN and CENP-HIKM sub-complexes (see Figure 6) (Klare et al., 2015a; Nagpal et al., 2015b). The difference in this region of CENP-C may be linked to the specific evolutionary path of *Drosophila*, which led to the loss of most CCAN subunits. DmCENP-C also contains two predicted AT-hooks domains (AT1 and AT2), which cannot be detected in the HsCENP-C and may mediate interactions with DNA (Heeger et al., 2005b). At the C-terminal region of DmCENP-C, there is a CENP-C motif for CENP-A binding and also a conserved dimerization domain (Heeger et al., 2005b; Sugimoto et al., 1997b).

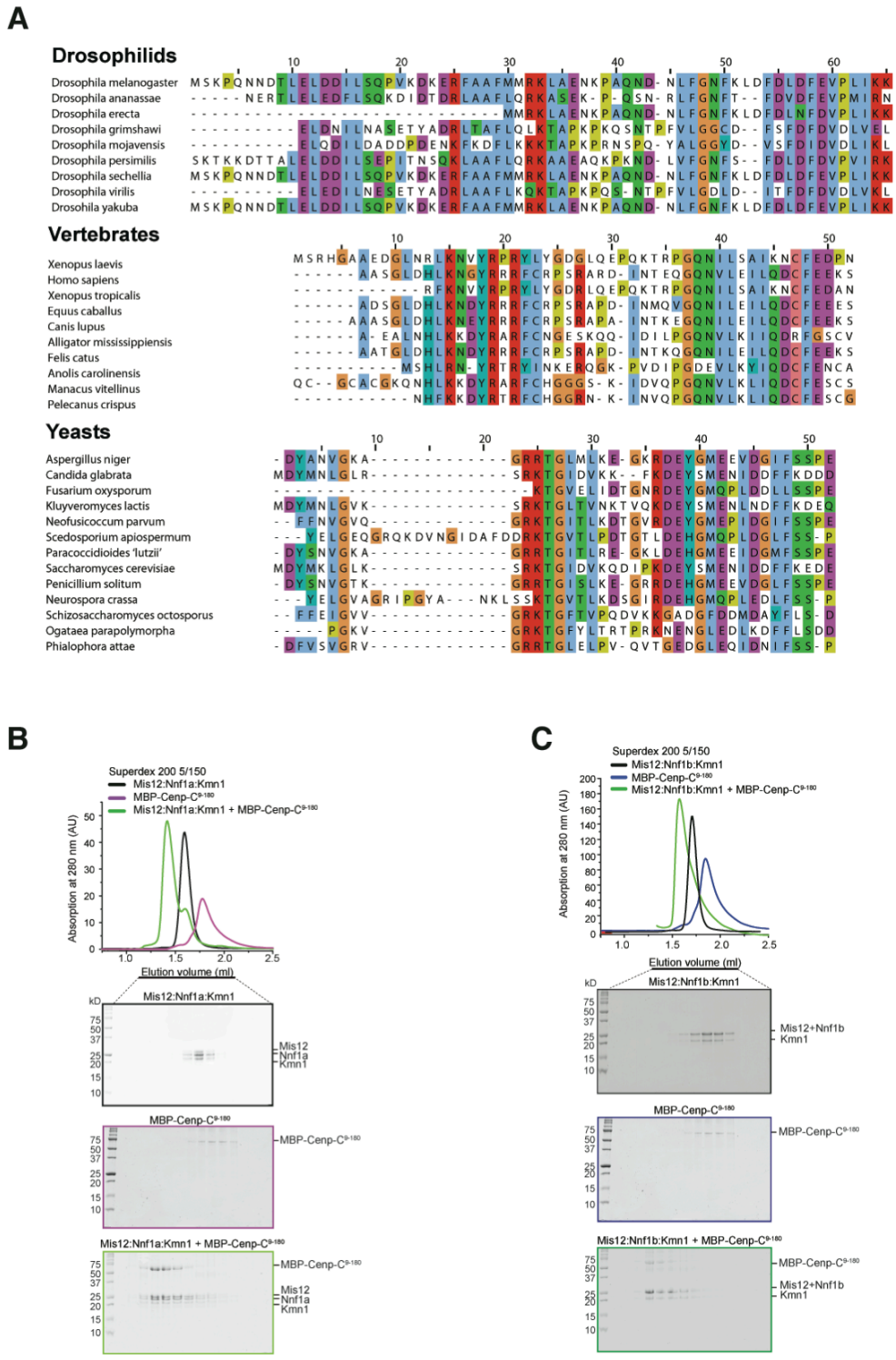

Figure 22 The DmMis12 complex binds the N-terminal region (without the first 8 amino acids) of CENP-C

(**A**) Alignments of N-terminal regions of CENP-C. Initial alignments of CENP-C in Drosophilids, vertebrates, and yeasts were created with PSI BLAST (http://blast.ncbi.nlm.nih.gov). A subset of sequences was then imported into the program MUSCLE (Edgar, 2004) for a refined alignment and visualized in Jalview (Waterhouse et al., 2009); Analytical size exclusion chromatography shows that the DmMis12a complex (**B**) and DmMis12b complex (**C**) binds directly to MBP-CENP-C$^{9-180}$.

Previous studies have demonstrated that Mis12-C binds directly to CENP-C (Hornung et al., 2014a; Przewloka et al., 2011a; Screpanti et al., 2011b). In humans, about 20 residues at the N-terminus of CENP-C are sufficient to bind with Mis12C (Screpanti et al., 2011b). In *S. cerevisiae* this is similar (Hornung et al., 2014a). Primary sequence alignment of the N-terminal region of CENP-C in different species did not exhibit strictly conserved features, but an overall pattern, such as a stretch of positive charges followed by hydrophobic amino acids, may somehow be conserved (see Figure 22A).

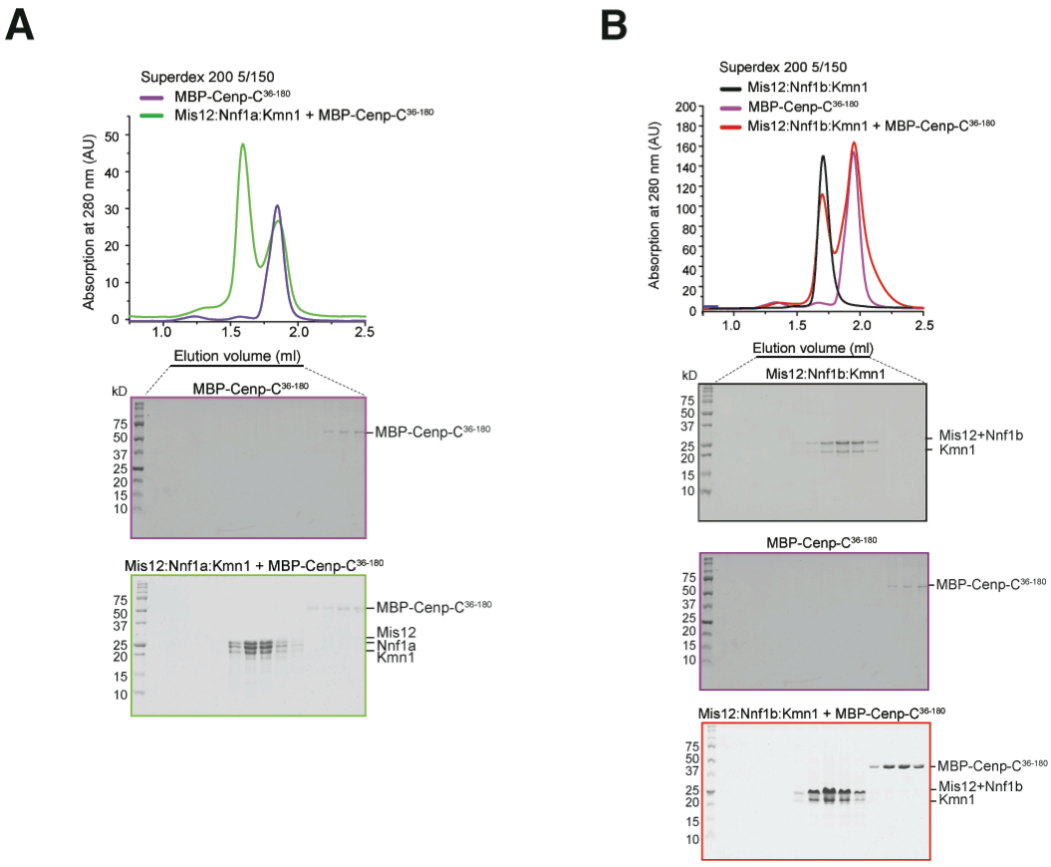

Figure 23 Important fragments for CENP-C binding to DmMis12 complexes

Size-exclusion chromatography analyses of deletion mutants of the N terminal CENP-C indicate that the first 35 amino acids are essential for the binding.

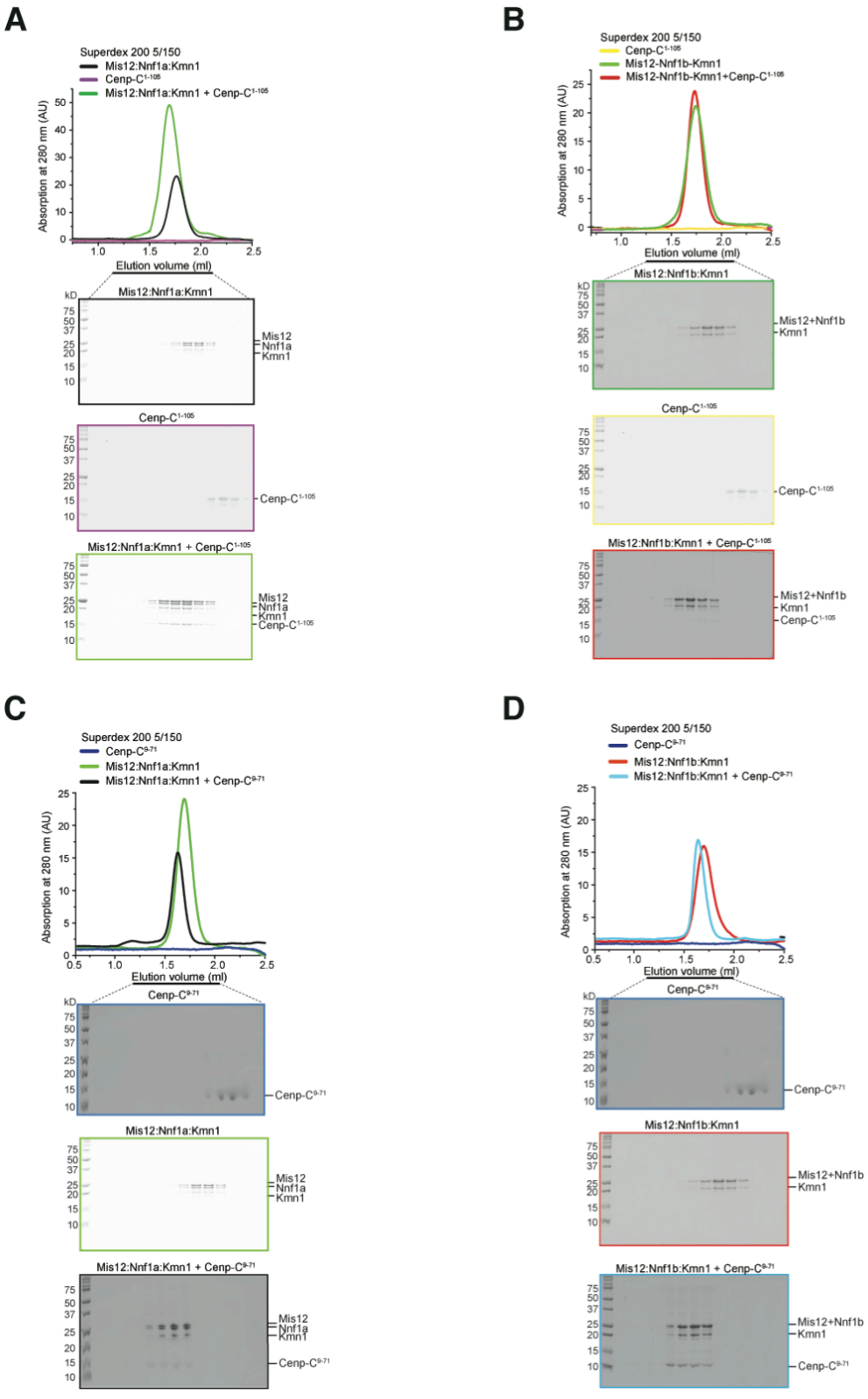

Figure 24 Minimal stable fragments of CENP-C bind to DmMis12 complexes

(**A**) and (**C**) suggest that different fragments of the N-terminal CENP-C can bind to the DmMis12a complex; In our experiments, the minimal stable fragment was CENP-C$^{9\text{-}71}$; Similar results were also obtained in the DmMis12b complex, which are shown in (**B**) and (**D**).

As we summarized at the beginning of this chapter, the DmCENP-C was shown to bind to DmMis12C in vivo (Przewloka et al., 2011b). However, the requirements for this binding interaction have not yet been mapped in detail. In order to test the binding, we expressed and purified a fusion protein of maltose binding protein (MBP) with a CENP-C fragment from 9 to 180 (CENP-C^{9-180}), and then incubated it with DmMis12C following the SEC experiment (see Figure 22B). Clear shifts of the elution profile to the left indicated that there is binding between these two samples. Identical results were observed for the DmMis12b complex (see Figure 22C). Sequence alignment indicated that residues 1-8 of DmCENP-C are not conserved in other drosophilids, but conservation increases significantly after this fragment. This implied that these 8 amino acids may not be involved in binding with CENP-C (see Figure 22A). However, larger N-terminal deletions (like DmCENP-C^{36-180}) inhibit the interaction with both the Mis12a and Mis12b complexes (see Figures 23A-B), indicating that residues 9-35 contain essential CENP-C binding elements.

The same strategy was used to test the effects of C-terminal deletions from the DmCENP-C N-terminal section. DmCENP-C^{1-105} is stable alone and interaction tests with the DmMis12a and DmMis12b complexes showed clear binding (see Figure 24A-B). An even shorter deletion mutant, DmCENP-C^{9-71} also binds to DmMis12C (see Figure 24C-D).

The sample of the DmMis12a: DmCENP-C^{1-105} complex in SEC experiments was monodisperse (see Figure 24). Sedimentation velocity experiments (Figure 25A) showed that the molecular mass of this complex is 76.4 (see Table 2), suggesting that the stoichiometry of DmMis12a:CENP-C is 1:1. Identical results were obtained with the Mis12b:CENP-C complex (see Figure 25B and Table 2). Further XL-MS experiments confirmed that CENP-C^{1-105} interacts with the Mis12 subunits or even Nnf1 and Kmn1 (see Figure 25C-D). Collectively, our biochemistry experiments demonstrate that despite

the fact that the sequence identity is poor at the N terminal end of CENP-C in different species, the DmMis12C complex binds directly to the N-terminal of CENP-C. This is consistent with results obtained for humans and yeast (Hornung et al., 2014a; Screpanti et al., 2011b).

This section has focused on the interaction between DmMis12C and CENP-C. As in humans, the N terminal of CENP-C is involved in binding with the Mis12 complex in *Drosophila* and the stoichiometry of this binding is one to one. In our biochemistry study, the residues 9-35 of CENP-C contain essential interaction determinants. Furthermore, we also identified a binding motif on DmMi12C. This would be the major topic for the next section.

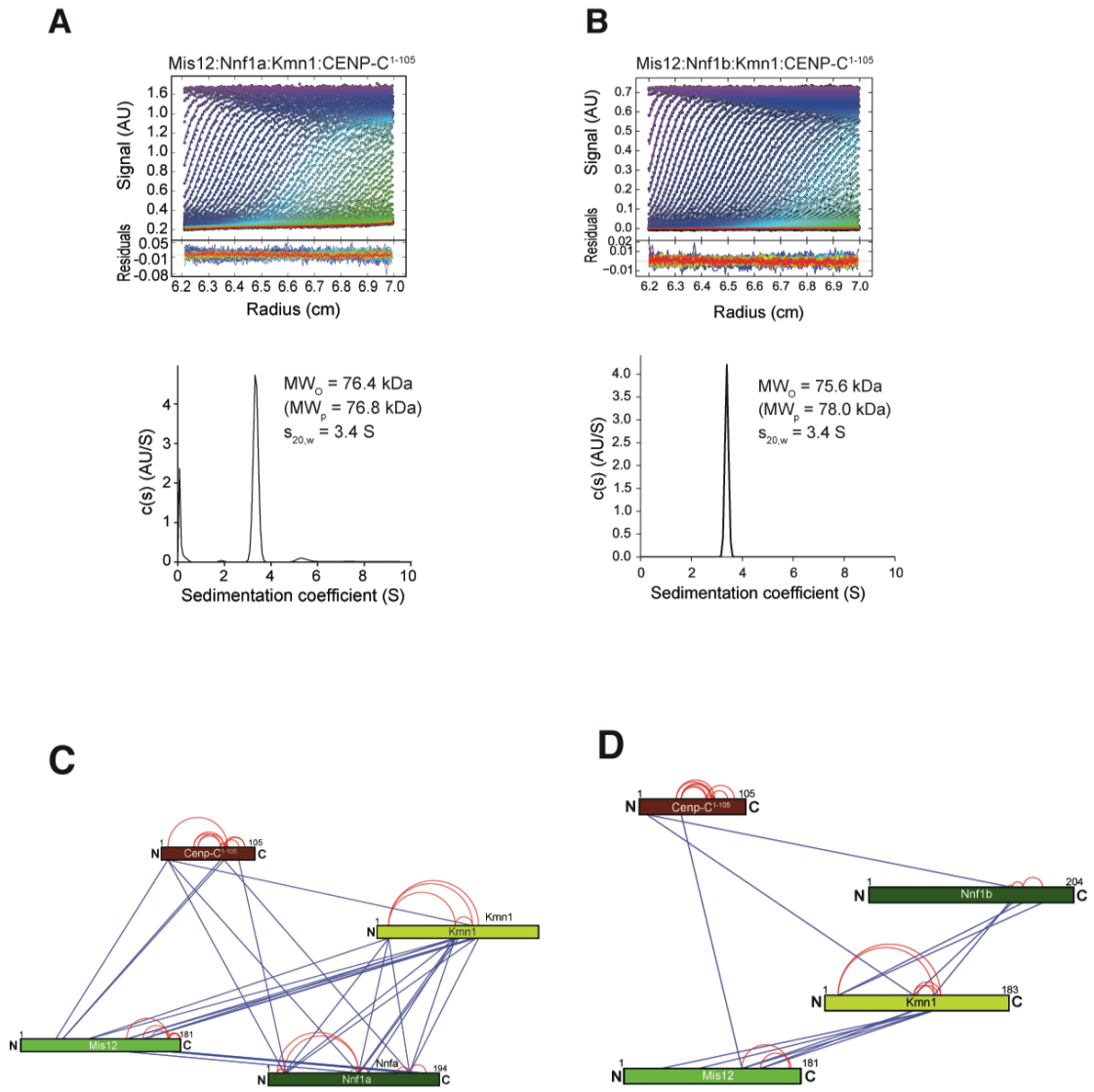

Figure 25 Biophysical analysis of the DmMis12-CENP-C$^{1-105}$ complex

Sedimentation velocity absorbance profiles of the DmMis12a (**A**) and DmMis12b (**B**) complexes with CENP-C$^{1-105}$, with residuals of fit showing the deviation of the c(*S*) model from the observed signals; the best-fit continuous-size c(*S*) distribution of the DmMis12a and DmMis12b complexes with CENP-C$^{1-105}$ is shown in the bottom section of the panel; Cross-linking-mass spectrometry (XL-MS) analysis of the DmMis12a (**C**) and DmMis12b (**D**) complex. Blue and red lines indicate inter- and intra-molecular cross-links, respectively.

3.4 A CENP-C binding site on the Mis12 subunit of the Mis12 complex

The sequence of the N-terminal region of CENP-C appears to be divergent in different species (see Figure 22A). However, both my previous data and others studies indicate that the function of this region as a Mis12 complex binding interface is conserved (Hornung et al., 2014b; Przewloka et al., 2007; Screpanti et al., 2011a). Previous biochemical experiments indicated that alanine substitutions of K10 and Y13 near the N-terminus of CENP-C reduce the interaction between CENP-C and the Mis12 complex in humans (Screpanti et al., 2011a). However, these amino acids are not conserved in Drosophila (see Figure 22A). Moreover, although Nnf1 has been shown to be important for DmMis12C binding to CENP-C *in vivo* (Przewloka et al., 2011b), no additional details of the motif in the DmMis12C essential for interaction with CENP-C were known at the start of this project.

A second structural prediction of Mis12 in DmMis12C suggests that there is a small helix at the very N-terminus (see Figure 19). This helix contains a stretch of hydrophobic amino acids, such as Alanine, Leucine, Phenylalanine and Tyrosine. In our attempts to crystallize the *Drosophila* Mis12 complex, we generated a version of the DmMis12a complex in which the first 15 residues of the Mis12 subunits had been deleted, referred to as the Mis12$^{\Delta N15}$ complex (see Figure 26). In contrast to the native DmMis12 complex (see Figure 12), the size exclusion chromatography experiment indicated that the Mis12$^{\Delta N15}$ complex contained two well-separated peaks (Peak 1 and Peak 2) (see Figure 26). The vertical black dashed bar is a reference indicating the elution volume of the native DmMis12 complex under the same conditions. Since Peak 2, indicated by the green arrow, is so close to the elution volume of the native DmMis12

complex, I concluded this peak should be a heterotrimer, similar to the native DmMis12 complex.

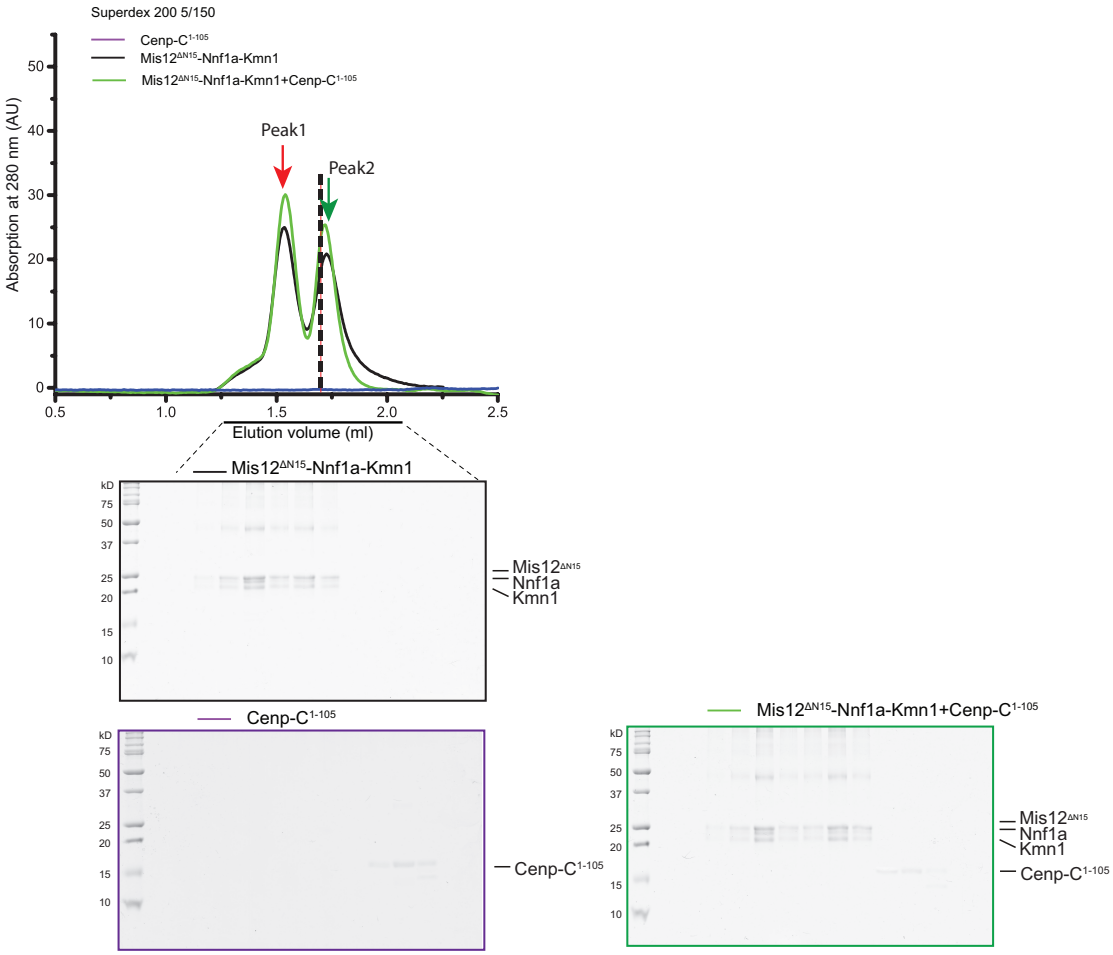

Figure 26 Additional size-exclusion chromatography analyses of mutant Mis12 complexes

Size-exclusion chromatography profile of a mutant of the DmMis12a complex in which 15 residues at the N-terminus of the Mis12 subunit were deleted (indicated as Mis12$^{\Delta N15}$). Note the complete lack of interaction with CENP-C$^{1-105}$.

However, Peak 1, indicated by the red arrow, eluted considerably earlier than Peak 2 (see Figure 26), while the SDS-PAGE gel did not reveal any difference in the gradients of these two peaks (see Figure 26). It is thus conceivable that Peak 1 is the result of the oligomerization of the heterotrimer in Peak 2.

I attempted to test the binding for the Mis12$^{\Delta N15}$ complex with the N-terminal of CENP-C, CENP-C $^{1-105}$. Unexpectedly, unlike the native DmMis12 complex depicted in Figure 24A, the deletion mutant failed to bind CENP-C $^{1-105}$ (see Figure 26). This implies that the first 15 amino acids in Mis12 are important for binding with the N-terminal of CENP-C, but the precise details of this effect remain unclear when I was doing the project.

Sequence alignment indicated that the first 15 residues of the Mis12 subunit are evolutionarily conserved (see Figure 27A). We hypothesized that this conservation may link with the binding of CENP-C and tested the role of three conserved phenylalanine (F) residues in this segment, F12, F13, and F15.

According to the SEC elution profile (see Figure 27B), a DmMis12a complex containing mutations F12D, F13D, and F15D in the Mis12 (indicated as Mis12$^{F12D, F13D, F15D}$) subunit was stable and monodisperse. As with the previous mutant lacking the first 15 residues of Mis12, this mutation lost the ability to interact with CENP-C$^{1-105}$ in a SEC experiment, indicating that the mutations prevent the interaction of Mis12 with CENP-C (see Figure 27B). Corroborating previous experiments, identical results were obtained with a DmMis12b complex with mutations of F12D, F13D, and F15D (see Figure 27C). Thus, our results demonstrate that the N-terminal region of the Mis12 subunit is crucial for the interaction of the DmMis12C with CENP-C.

This section demonstrates that the N-terminal region of the Mis12 subunit in the Mis12 complex is also very important for binding CENP-C. Furthermore, it reveals that three very conserved phenylalanine residues (F12, F13, F15) in this region of Mis12 are a necessary determinant in Mis12 for binding with CENP-C *in vitro*.

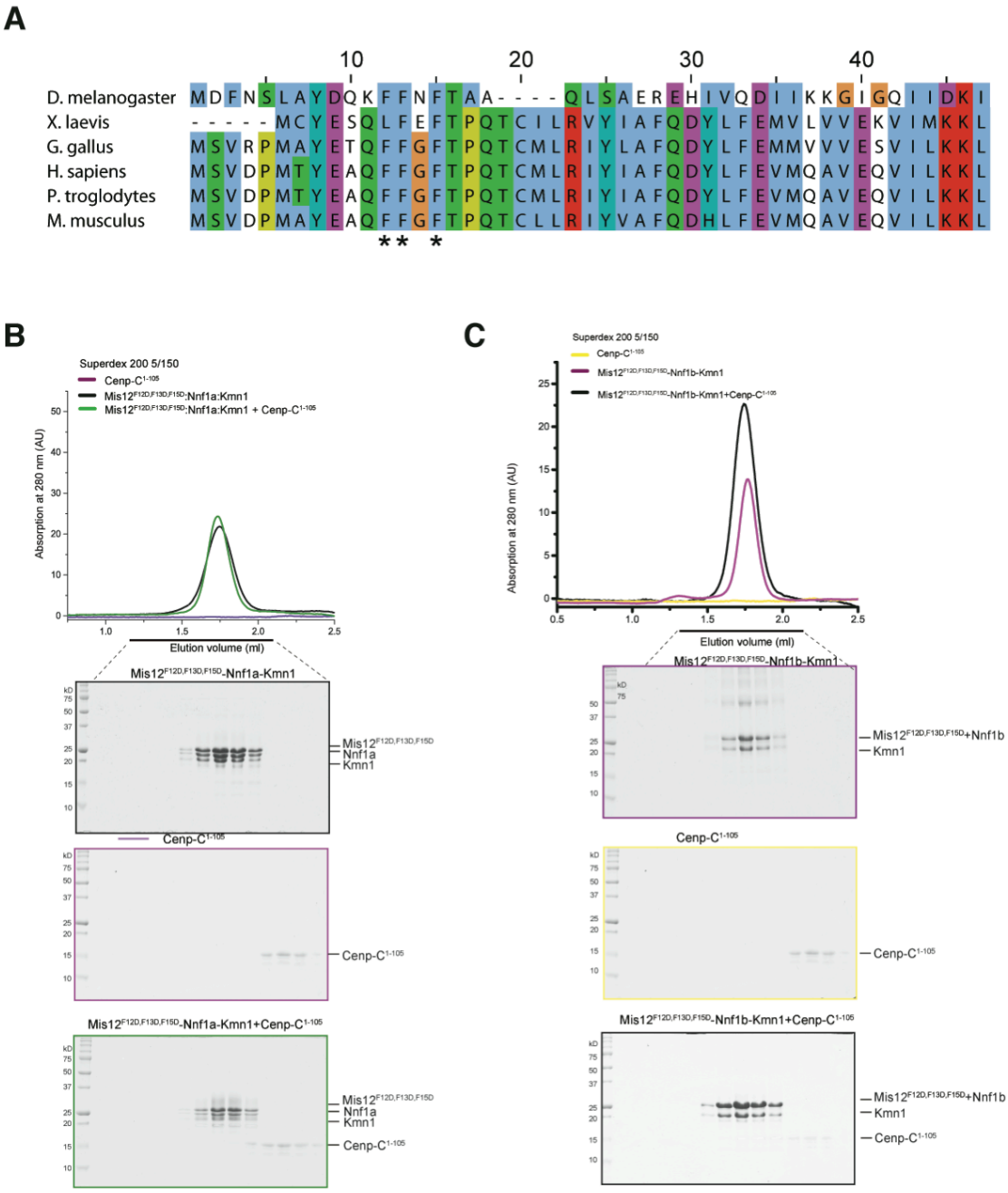

Figure 27 A CENP-C binding region in the Mis12 subunit

(**A**) Sequence alignment of the N-terminal region of the Mis12 subunit of the Mis12 complex. The position of three phenylalanine (F) residues that were mutated to Asp are indicated by asterisks; (**B**) Analytical size exclusion chromatography shows that the CENP-C$^{1\text{-}105}$ is unable to interact with the mutant DmMis12a complex; (**C**) Similar results were obtained in the DmMis12b complex.

3.5 Reconstitution of other KMN network subunits

In the previous sections, we showed that despite the fact that one subunit is absent, the DmMis12 complexes bind to the N terminal of CENP-C, meaning that Mis12 complexes in *Drosophila* use a similar method to humans and yeast to link to the inner kinetochore. However, we also wished to determine how the Mis12 complex in Drosophila interacts with the other subunits of the KMN network: the Knl1 complex and the Ndc80 complex. We first investigated Knl1/Spc105R in *Drosophila*.

The Knl1/Spc105R complex in *Drosophila* differs from that of other eukaryotes in that one of the two subunits of the complex, Zwint, is missing in *Drosophila* (see Figure 28A). Moreover, DmSpc105R (Knl1 homology in *Drosophila*) is shorter in Drosophila than in humans (see Figure 28A).

It has previously been shown that the high-resolution crystal structure of the C-terminal region of human Knl1/Spc105R contains two consecutive RWD (RING finger, WD repeat, DEAD-like helicases) domains, preceded by a coiled-coil region (Petrovic et al., 2014a). The RWD domain mediates binding to the C terminal region of the Nsl1 in humans, and the coiled-coil region mediates the interaction with Zwint, which also contains coiled-coil parts (Kiyomitsu et al., 2010; Petrovic et al., 2014b; Petrovic et al., 2010a) (see Figure 28A).

DmSpc105R appears to diverge from this structure. First, unlike the ~200-residue coiled-coil domain identified in the C terminal of human Knl1, the program COILS (Lupas et al., 1991) failed to identify an equivalently extended coiled-coil region within the C-terminal region of DmSpc05R. Second, it remains unclear whether the C-terminal region of DmSpc105R contains the RWD domains identified in its human counterpart

Knl1. BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) searches with the last ~200 residues of the C-terminal region of DmSpc105R failed to detect homologous proteins beyond drosophilids. However, other software based on the profile-profile alignment and fold recognition algorithm, such as FFAS (Xu et al., 2014), readily identified the C-terminal region of human Knl1, which contains two RWD domains, as a possible structural homologue for the C-terminal region of *Drosophila* Knl1/Spc105R.

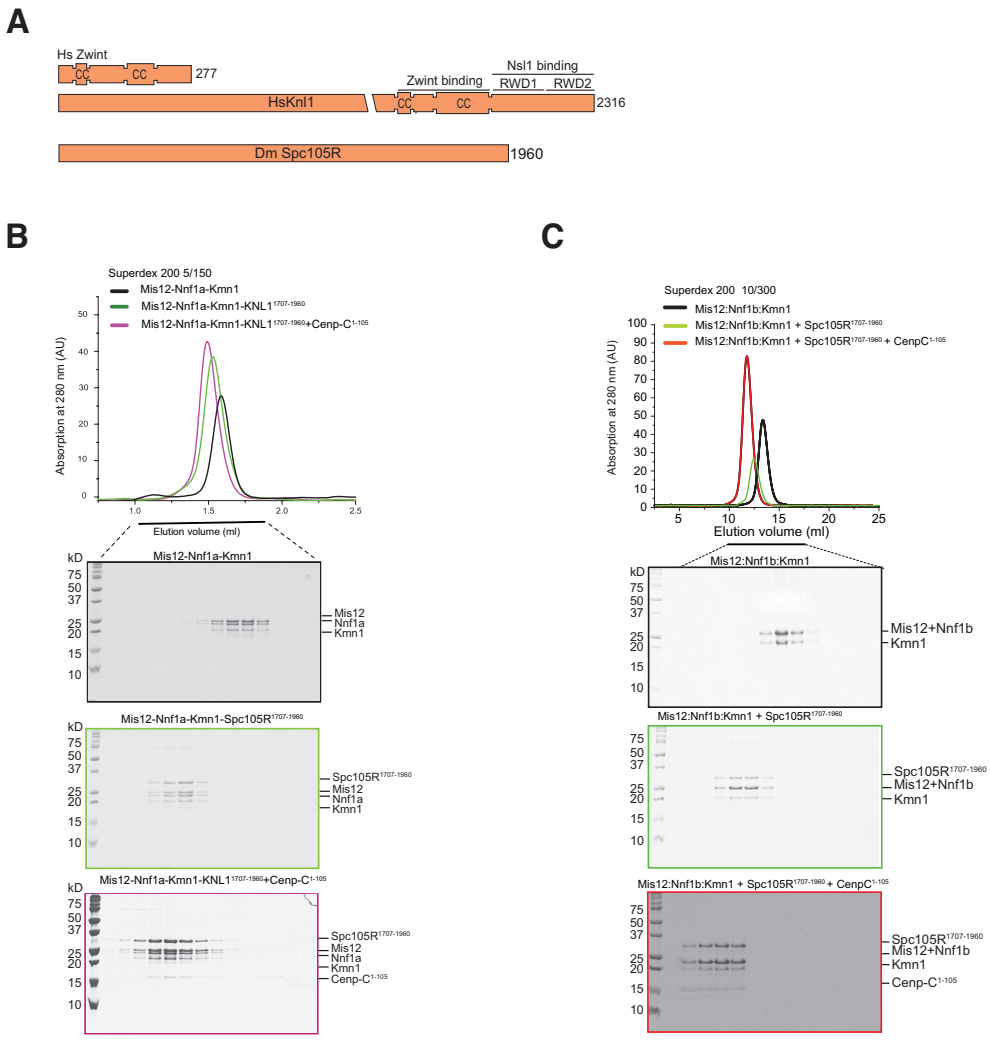

Figure 28 Interaction of the Mis12 complex with the C-terminal region of DmSpc105

(**A**) Schematic comparison of the domain structure of Spc105 in *Drosophila* and of its human homologue Knl1/CASC5. The C-terminal region of Knl1/Spc105R contains a coiled-coil domain that has been implicated in a direct interaction with Zwint, a coiled-coil protein that has not been identified in *Drosophila*. It also contains two consecutive RWD domains implicated in a direct interaction with the Mis12 subunit

Nsl1 (Petrovic et al., 2014b); (**B**) size-exclusion chromatography analysis of the indicated complexes demonstrates that Spc105R1$^{707-1960}$ and CENP-C$^{1-105}$ enter a single complex with the Mis12a complex; (**C**) The Mis12b: Spc105R$^{1707-1960}$ complex binds CENP-C$^{1-105}$.

Sequence alignment of the last 230 amino acids indicated that the sequence conservation was poor for *Drosophila* in comparison to the other species (see Figure 29A). Conservation was limited to hydrophobic residues (shown in blue in Figure 29A), probably reflecting conservation of secondary structure. We tried to compare the predicted secondary structure of this section in *Drosophila* to the structures in humans, for which a crystal structure is already available (Petrovic et al., 2014a). The secondary structure predictions of the C-terminal region of DmSpc105R reveal similarities between *Drosophila* and the human homologue, as shown in Figure 29B. This analysis suggests that the C-terminal region of DmSpc105R also contains two RWD domains.

Despite this progress, there is still much to be determined regarding the structural organization of the C-terminal region of DmSpc105R. Previous studies have demonstrated that a ~600-residue construct containing the C-terminal region of DmSpc105R can interact with Mis12 complex in a yeast two-hybrid (Y2H) experiment (Schittenhelm et al., 2009b). This implies that the C-terminal regions of the human and fly sequences should be functionally related.

To clarify the interaction of DmSpc105R with the Mis12 complex, we first expressed and purified different kinds of DmSpc105R fragments. Constructs with or without the predicted coiled coil region (the segment 1852-1889) proved to be insoluble (See Table 6). Co-expression strategies of Mis12a or Mis12b complexes with different segments containing the C-terminal region of Knl1/Spc105R were also included in the expression tests (See Table 6). In these tests, only Knl1/Spc105R$^{1707-1960}$ could be co-expressed in a soluble form with the Mis12a and Mis12b complexes. SEC studies indicated that, in both cases, an apparently monodisperse and stoichiometric complex

was formed (see Figure 28B-C). According to AUC sedimentation velocity experiments, the predicted stoichiometry of DmMis12C: Knl1/Spc105R is 1:1 (Table 2 and Figure 30A-B).

Our biochemical experiments were not able to identify the binding details for the C-terminal of *Drosophila* Knl1/Spc105R to the Mis12 complex. Taking the structural conservation of both *Drosophila* Knl1/Spc105R and Mis12 complexes from the human counterparts into account, we hypothesize that the C terminal region of Knl1/Spc105R binds to the thinner end of the Mis12 complexes in *Drosophila*. However, some differences between the C-terminal region of Knl1/Spc105R in *Drosophila* and its human homolog are evident. In humans, this double RWD domain was quite stable and it could be purified and crystallized (Petrovic et al., 2014a). However, in *Drosophila* extensive screens of the constructs of the C-terminal region of Knl1/Spc105R have not identified any stable fragment that can be expressed and purified in isolation. This domain was only stable when co-expressed with the DmMis12 complex, which may imply that the binding details of the C-terminal region of Knl1/Spc105R in *Drosophila* to the Mis12 complex differ in from those of the human proteins.

Thus, we tried to reconstitute two sub-complexes of KMN network, Mis12 complex and Knl1. We also tried to extend our reconstitution to Ndc80 complex. In both humans and yeast, the thinner end of the Mis12 complex also connects to the RWD domains in the Ndc80 complex (Dimitrova et al., 2016; Malvezzi et al., 2013; Petrovic et al., 2016; Petrovic et al., 2010b). The C-terminal regions of the Nsl1 and Dns1 subunits have been shown to be important for the recruitment of the Ndc80 complex in yeast and humans (Malvezzi et al., 2013; Petrovic et al., 2010b). Since one of these subunits is missing in the Mis12 complexes of *Drosophila*, it is not clear whether binding to Ndc80 in *Drosophila* is through a conserved mechanism. In *C. elegans*, which also lost most of the CCAN subunits, the Knl1/Spc105R and the Mis12 complex form a joint binding platform for

the Ndc80 complex (Cheeseman et al., 2006). In *Drosophila*, the C-terminus of Knl1/Spc105R remains on kinetochores localized at the junction of the DmMis12 and DmNdc80 complexes (Venkei et al., 2012b). This may suggest that Knl1/Spc105R is important for stabilization of the Ndc80 complex in both *Drosophila* and *C. elegans*. We tried to express and purify DmNdc80 complex, but to date our efforts have been unsuccessful. The next step will be co-express and purify the entire KMN network in *Drosophila* and to endeavour to clarify the binding details at this thinner end of the Mis12 complexes.

SEC experiments indicate that both the Mis12a:Knl1/Spc105R$^{1707\text{-}1960}$ and the Mis12b:Knl1/Spc105R$^{1707\text{-}1960}$ complexes can also incorporate CENP-C$^{1\text{-}105}$ in single monodisperse complexes (see Figure 28B-C). Further AUC sedimentation velocity analysis of the Mis12a:Knl1/Spc105R$^{1707\text{-}1960}$:CENP-C $^{1\text{-}105}$ complex revealed a stable 1:1:1 assembly (see Table 2 and Figure 30C).

Table 6 Solubility test of the C terminal of Knl1/Spc105R

Expression ways	Constructs	Solubility
Single expression	Knl1/Spc105R$^{1707\text{-}1882}$	No
	Knl1/Spc105R$^{1707\text{-}1960}$	No
	Knl1/Spc105R$^{1875\text{-}1960}$	No
	Knl1/Spc105R$^{1887\text{-}1960}$	No
	Knl1/Spc105R$^{1890\text{-}1960}$	No
Co-expression	MAK/MBK-Knl1/Spc105R$^{1707\text{-}1882}$	No
	MAK/MBK- Knl1/Spc105R$^{1707\text{-}1960}$	Yes
	MAK/MBK- Knl1/Spc105R$^{1875\text{-}1960}$	No
	MAK/MBK- Knl1/Spc105R$^{1887\text{-}1960}$	No
	MAK/MBK- Knl1/Spc105R$^{1890\text{-}1960}$	No

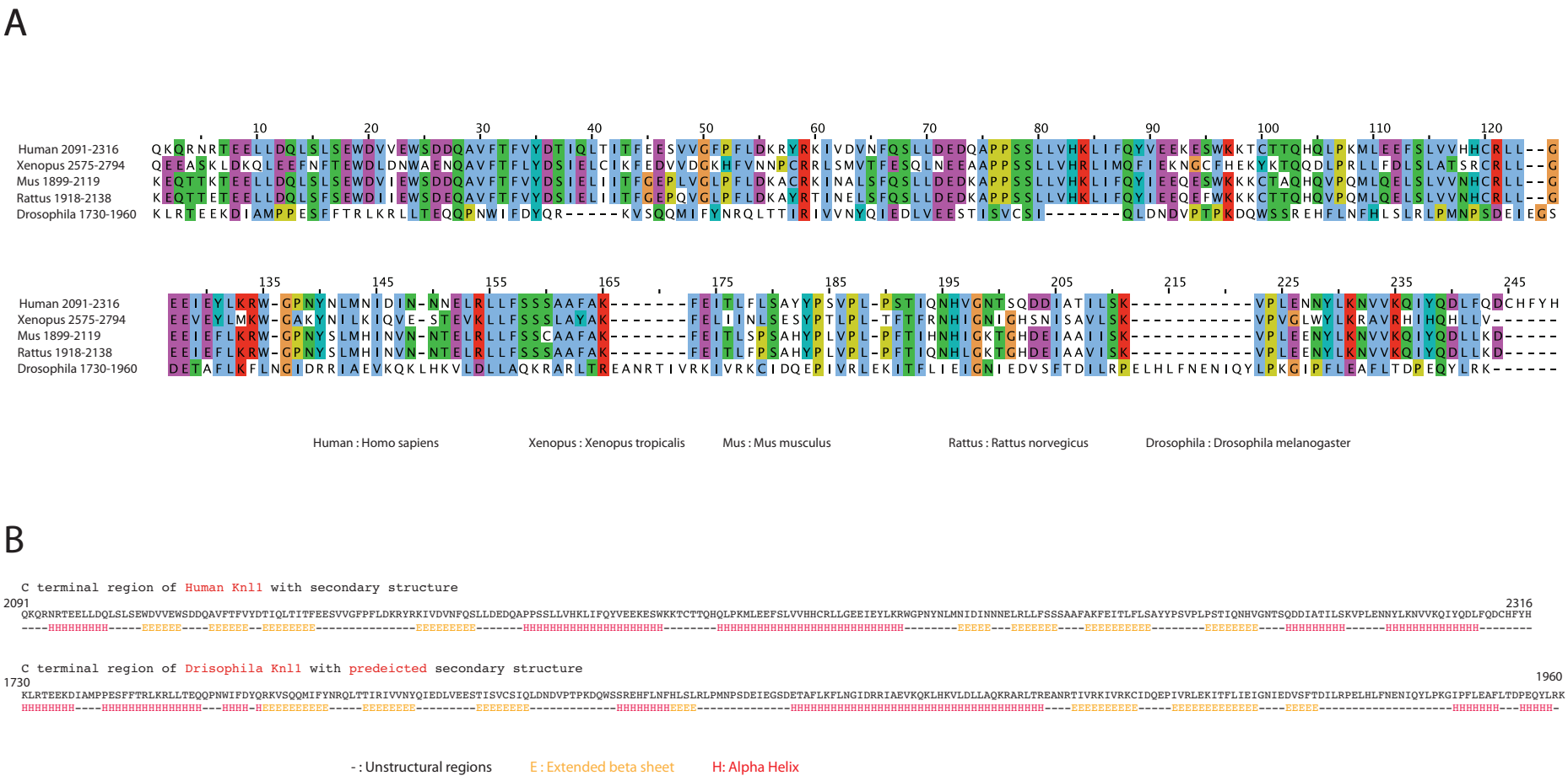

Figure 29 Sequence alignment and secondary structure comparison

(**A**) Multiple sequence alignment of the C-terminal regions of Knl1/Spc105R from the indicated species. The alignment was calculated with the program Muscle (http://www.ebi.ac.uk/Tools/msa/muscle/); (**B**) Second structural comparison of the C-terminal Knl1/Spc105R from humans with predictions for Drosophila.

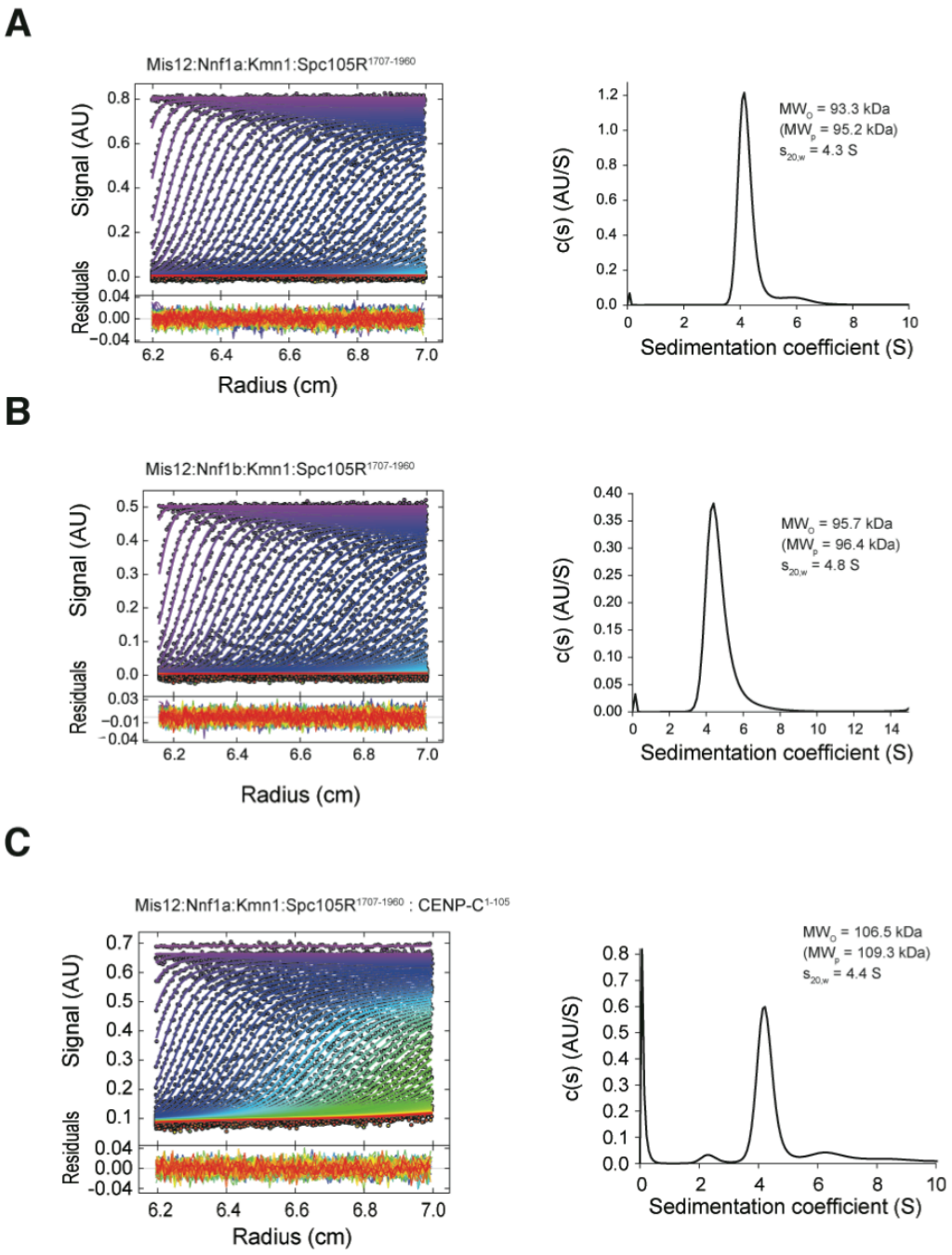

Figure 30 Analytical ultracentrifugation analyses of the complexes of DmMis12a and DmMis12b with Spc105R

Sedimentation velocity absorbance profiles of the DmMis12a and DmMis12b complexes with Spc105R1707-1960, with residuals of the fit showing the deviation of the c(S) model from the observed signals; the best-fit continuous-size c(S) distribution of the DmMis12a:Knl1/Spc105R1707-1960 (A) DmMis12b:Knl1/Spc105R1707-1960 (B) , and DmMis12a:Knl1/Spc105R1707-1960:CENP-C1-105 (C) Complexes are shown on the right-hand side of the panel.

3.6 The structure of Mis12C and details of the binding with CENP-C in humans

In our attempts to crystallize the Drosophila Mis12 complex, we generated DmMis12a and DmMis12b complexes and different kinds of deletion mutations. However, to date crystallization tests have been unsuccessful. Together with Dr. Arsen Petrovic and Dr. Jenny Keller, we obtained a high-resolution structure of the human Mis12 complex (Mis12C) associated with the N-terminal region of CENP-C (CENP-C 1-71) (see Figure 31A) (Petrovic et al., 2016). At the same time, Prof. Dr. Stephen C. Harrison's group also obtained the crystal structure of the MIND complex (Mis12C homology in yeast) in complex with the N-terminal region of Mif2 (CENP-C homolog) in yeast (see Figure 31B) (Dimitrova et al., 2016).

3.6.1 The structure of Mis12C and the N terminal of CENP-C

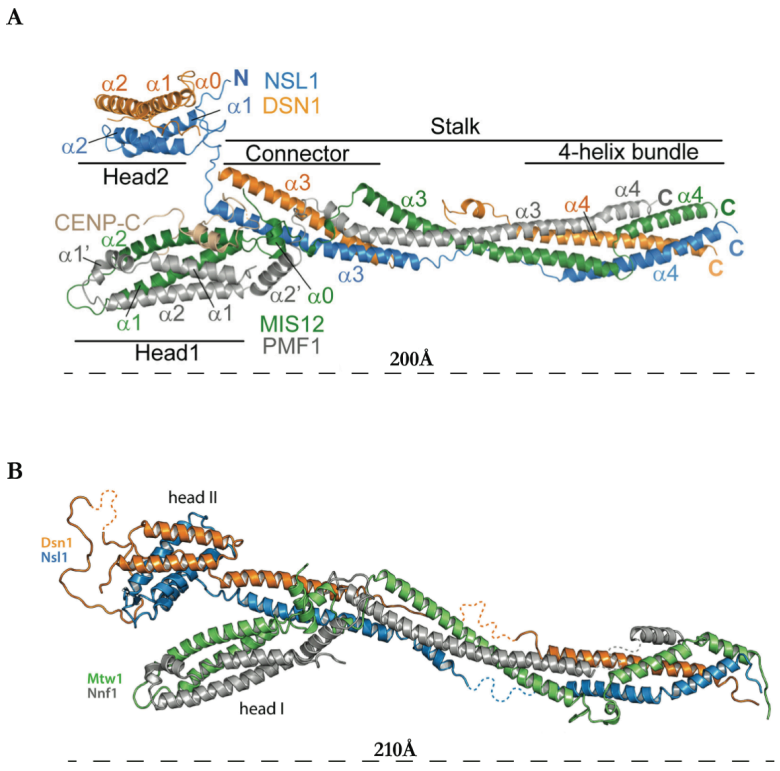

Figure 31 Structure of Mis12 complex in human and yeast

(**A**) Cartoon diagrams of the Mis12C^{nano}:CENP-C$^{1-71}$ complex. All the alpha helices have been labeled with α followed by a number. Figure adapted from (Petrovic et al., 2016); (**B**) Cartoon representation of the crystal structure of *Kluyveromyces lactis* MIND complex. Figure adapted from (Dimitrova et al., 2016) .

Our results provide structural organization details for all four subunits of Mis12C in humans. Consistent with negative stain EM structure (see Figure 14B), the crystal structure of Mis12C in humans shows that it is an elongated rod and the long axis of the complex is approximately 200 Å (Petrovic et al., 2016).

As shown in Figure 31A, the four subunits of the Mis12 complex cross the whole length of the complex and are co-linear, going from N terminus to C terminus. The complex is composed of two distinct sub-complexes, Mis12:Pmf1 and Nsl1:Dsn1. The

N terminus of the Mis12C consists of two globular heads. Head 1 contains the N-termini of Mis12 and Pmf1, whereas the N-terminal regions of Dsn1 and Nsl1 constitute Head2. It is interesting that the N-terminal regions of all four subunits of Mis12C contain a helical hairpin structure. In both Head 1 and Head 2, two helical hairpins from different subunits pack together and form four-helix bundles structures. Head 1 is less flexible and more firmly attached to the coiled coil stalk that we designated as "connector", which come from Dsn1:Ns1l sub-complex. However, Head 2 seems to be detached and more flexible than Head1. The long Mis12:Pmf1 coiled coil that forms the binding site for the "connector" is the core of the structure. At the very C terminal of Mis12C, all subunits meet to form 4-stranded helical bundle.

Although specific sequence similarity between humans and yeast is limited, the overall structure of Mis12C in yeast is highly similar to the Mis12 complex in humans (see Figure 31B) (Dimitrova et al., 2016).

Mis12C binds to the N-terminal segment of CENP-C, and this binding is highly conserved in different species (Liu et al., 2016; Richter et al., 2016; Screpanti et al., 2011a). The high-resolution structure of our human Mis12C also reveals the binding details with the N-terminal of CENP-C.

In general, the helical hairpins in Head 1, together with the helical connector in Mis12C offer the platform for binding the N-terminal segment of CENP-C(see Figure 31A). CENP-C forms an asymmetric "U", which is unstructured in most regions, and rides on this platform. Residues 6 to 22 bind to the helical hairpin formed by the α1 and α2 alpha helices of Mis12 in Head 1. Residues 28 to 30 form the bottom of the asymmetric "U", followed by a small helix (residues 32 to 44). Unfortunately, the remains of the CENP-C fragment is not visible in our structure. This is in all likelihood due to the fact that it is not involved in interaction and therefore flexible. Previous research has indicated that residues Lys10 and Tyr13 in the N-terminal of CENP-C are necessary for

tight binding of CENP-C to Mis12C, and this was explained by the structure we obtained (Petrovic et al., 2016; Screpanti et al., 2011a). Our research also indicates that the first small alpha helix (α0), especially residues Tyr8, Phe12, and Phe13, in Mis12 are important for the binding of Mis12C to CENP-C. However, our high-resolution structure doesn't show direct binding between α0 and CENP-C, which suggests that this alpha helix may be involved in stabilizing the interaction of Mis12C with CENP-C indirectly, by reinforcing a particular conformation of the complex (Petrovic et al., 2016).

3.6.2 The mechanism of Aurora B mediates CENP-C binding to Mis12C

Previous studies indicated that the Mis12 complex appears in the nucleus at interphase, but can only bind to CENP-C at centromeres during mitosis (Rago et al., 2015b). Two residues in the Dsn1 subunit of the Mis12 complex, Ser100 and Ser109, are substrates of the mitotic kinase Aurora B and the phosphorylation of Dsn1 increases the binding affinity between Mis12 complex and CENP-C (Kim and Yu, 2015). Ser100 and Ser109 localize in an unstructured region located N-terminally to Head 2 of Dsn1 and that was not included in our crystallized construct. The deletion mutant indicates that the flexible region and Head 2 are not essential for the binding with CENP-C. Interestingly, deletion of the region preceding Head2 causes Mis12C to bind more tightly to CENP-C (60-fold more). This effect can be recapitulated with shorter deletions of 10 residues encompassing Ser100 and Ser109.This implies that the fragment containing Ser100 and Ser109 in Dsn1 interferes with the binding to CENP-C, and that phosphorylation of

these two amino acids by Aurora B releases this inhibition, and promotes Mis12C binding to CENP-C (see Figure 32B).

This hypothesis is also confirmed by the sequence alignment, which indicates that two motifs encompassing Ser100 and Ser109 in Dsn1 are similar to a fragment at the N-terminal of CENP-C (see Figure 32A). In addition to the effects observed with the deletion mutants, point mutations in positively charged residues in these Dsn1 motifs (Figure 32) also increased the binding affinity between Mis12C and the N-terminal of CENP-C (Petrovic et al., 2016). In biochemical experiments, we did not measure binding between segments of Dsn1 encompassing the Ser100, Ser109 and Head 2 and the rest of the Mis12C, suggesting that the ability of this region of Dsn1 to interfere with CENP-C binding relies on the high effective concentration of the Dsn1 fragments (Petrovic et al., 2016). As in humans, research from Prof. Dr. Stephen C. Harrison's group also indicates that phosphorylation of Head 2 by Aurora B releases Head 1 to bind to the N-terminal of Mif2 (CENP-C homology in yeast) (Dimitrova et al., 2016).

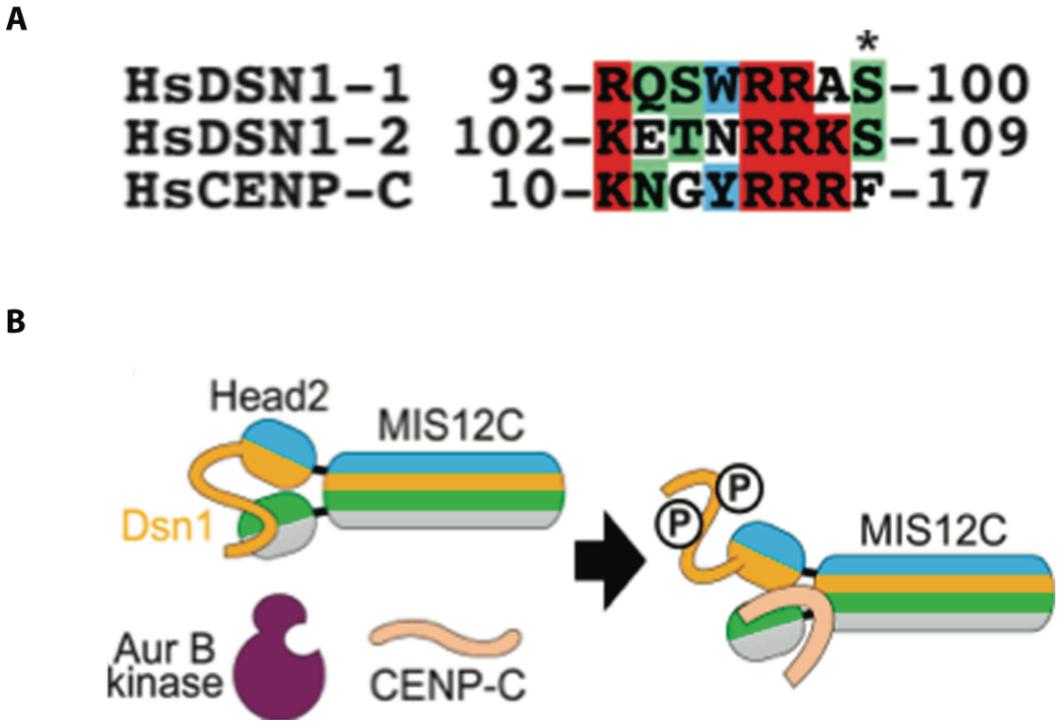

Figure 32 Aurora B regulates Mis12C binding to CENP-C

(**A**) Two similar motifs in Dsn1 at the Aurora B phosphorylation sites (Ser100 and Ser109) aligned with very N-terminal of CENP-C; (**B**) Diagram shows the mechanism of Aurora B's regulation function in Mis12C binding to CENP-C. From (Petrovic et al., 2016).

3.7 Structural and functional analysis of DmMis12C and its binding with CENP-C

In this section, I will analyze the way in which the high-resolution structures of Mis12C in humans and yeast contribute to our understanding of the structure and function of Mis12C in Drosophila.

3.7.1 Mis12 and Nnf1 form the backbone of DmMis12C

My biochemistry and negative stain EM results indicate that although one subunit is missing, the overall structure of DmMis12C remains similar to its homologs in humans and yeast (See Figure 14). Here, I further investigate the details of these structures.

As already reported, previous studies suggested that Mis12 complexes could be separated into two sub-complexes, Mis12:Pmf1 and Dsn1:Nsl1 (Maskell et al., 2010a; Petrovic et al., 2014a; Petrovic et al., 2010b). The structure of Mis12C in humans and yeast are also consistent with this hypothesis, and the interactions between Mis12 and Nnf1 extend from the N-terminus to the C-terminus (Dimitrova et al., 2016; Petrovic et al., 2016). Our biochemical results confirmed that this interaction takes place in *Drosophila* for Mis12 and either Nnf1a or Nnf1b (see Figure 9B). This binding appears to be well conserved and creates the binding platform for Kmn1 (Table 1 and Figure 12). Crosslinking combined with MS analysis revealed a very extensive network of interactions between the Mis12 and Nnf1a or Nnf1b subunits, extending along the length of their sequences in DmNnf1a and DmNnf1b complexes (See Figure 21B-C). This strongly implies that Mis12, Nnf1a and Nnf1b are not globular, but rather elongated in the complex. In agreement with this hypothesis, most of our deletion mutants on

Mis12 and Nnf1a/Nnf1b were unsuccessful (See Figure 19). The elongated Mis12-Nnf1a or Nnf1b interact with each other and impose the overall shape of the Mis12 complex in *Drosophila*, as seen in Figure 14. A loss of either Mis12 or Nnf1 destroyed the complex completely (see Table 1), strongly suggesting that Mis12 and Nnf1 form the backbone of the Mis12 complex in a manner similar to that in humans and in yeast (see Figure 31A-B).

Inside the backbone, the human structure indicates that two long alpha helixes from Mis12 and Pmf1 (two α3 see Figure 31A) stabilize and connect the N terminus to the C terminus. A similar structure appears to be conserved in both the DmMis12a and the DmMis12b complex, where the deletion mutants at both of these predicted coiled coil regions destroyed the complexes (see Table 4 and Table 5).

My results indicate that the backbone formed by Mis12 and Nnf1a or Nnf1b offers a platform for Kmn1 to bind (see Table 4 and Table 5). Secondary structural predictions, using different software, consistently indicated that the C-terminal half of Kmn1 contains long alpha helices. Since the overall structure of the DmMis12 complex is an elongated rod as in humans and yeast, it seems reasonable to assume that the C-terminal region of Kmn1 binds to the C-terminal regions of the Mis12:Nnf1a/Nnf1b sub-complex to form a three-helix bundle.

3.7.2 The interaction between DmMis12C and CENP-C

In my biochemical study, I discovered that both the DmMis12a and the DmMis12b complex can interact with the N-terminal of CENP-C, as has been shown in humans and yeast (Hornung et al., 2014b; Screpanti et al., 2011a). In a high-resolution structural study of Mis12C in humans and yeast, both Prof. Dr. Stephen C. Harrison's

group and our own group found that Head 2 regions are not required for CENP-C binding (Dimitrova et al., 2016; Petrovic et al., 2016). The biochemical studies from the next section indicated that in *Drosophila*, a similar head region is also not essential for CENP-C binding. CENP-C in *Drosophila*, humans and yeast are highly divergent in both length and sequence. However, this binding seems to be fairly conserved in most organisms.

In my deletion tests of the DmMis12C, I discovered the predicted small alpha helix at the N-terminal of Mis12 is important for DmMis12C binding to CENP-C (see Figure 26). Furthermore, I discovered that three-conserved Phenylalanine (F12, F13, F15 in *Drosophila*) is crucially important for this interaction. Our high-resolution structure of Mis12C in humans indicates that at the N-terminus of the Mis12 subunit, there is a small alpha helix α0 (see Figure 31A), and the mutations of the conserved amino acids ($Tyr8^{Mis12}$, $Phe12^{Mis12}$, and $Phe13^{Mis12}$) disrupts Mis12C binding to the N-terminal of CENP-C in humans. These mutations also inhibit Mis12C binding to CENP-C and CENP-T *in vivo* (Petrovic et al., 2016). I can thus conclude that the CENP-C binding motif for Mis12C is conserved in *Drosophila* and humans.

In general, the interface between the Mis12 complex and CENP-C is probably highly conserved in all eukaryotes.

3.7.3 Nsl1 could be the lost subunit in DmMis12C

My previous experiments indicated that there are two Mis12C in *Drosophila*, and unlike in humans and yeast, there are only three subunits in DmMis12C (see Figure 11 and Figure 12). However, although one subunit is missing, the overall structure of DmMis12C bears some similarity to its human and yeast homologs (see Figure 14). My

biochemical results indicate that the Mis12 binding ability of both CENP-C and Knl1 is well conserved in *Drosophila* and I hypothesize that the structure and function of Mis12C will also be well conserved in all eukaryotes. Consequently, it is of interest to determine which subunit is missing in *Drosophila* in order to understand the role of different subunits in Mis12C during mitosis.

Based on the sequence alignment, previous research has proposed that Dsn1 is the missing subunit in the *Drosophila* kinetochore (see Figure 32A) (Przewloka et al., 2007; Schittenhelm et al., 2007). However, for both Nsl1 and Dsn1, sequence alignment comparison indicates that sequence similarities are quite low in humans, *C. elegans*, and yeast (Schittenhelm et al., 2007), and to date no *in vitro* or *in vivo* studies support this proposition. In this section, I endeavor to gain some primary insights into the missing subunit in DmMis12C.

3.7.3.1 Head2 is structurally conserved in DmMis12C

Both human and yeast Mis12C contain a flexible Head 2 region (see Figure 31A-B). From the secondary structure prediction (see Figure 19) and the biochemical results (see Table 3 and Figure 20), the N-terminal region of Kmn1 ($Kmn1^{1-68}$, which I named as Head region in *Drosophila*) is a flexible section, whose deletion does not affect the stability of DmMis12 complex (see Figure 20). CD experiments indicated that approximately 65% of this head region is unstructured (see Figure 20C). I suppose that the Head region in Kmn1 is the counterpart of Head 2 in both humans and yeast. As in these two species, the Head region in Kmn1 is likely to be located at the thicker end of the rod particles and may contribute to the overall thickness of this section of the Mis12C. In the following sections, I will use "Head 2" as a synonym for the "head region" in Kmn1.

3.7.3.2 Functional analysis of Head2 in DmMis12C

The Head 2 in both humans and yeast contains the N-terminal regions of Nsl1 and Dsn1. Ser100 and Ser109 in the latter are substrates of Aurora B, which mediates the Mis12C localization to the centromere (see Figure 32B) (Dimitrova et al., 2016; Kim and Yu, 2015; Petrovic et al., 2016). I therefore resolved to determine whether the Head 2 in DmMis12C has a similar function to its human and yeast homologs.

In humans, it has been shown that Head2 is not required for CENP-C binding (Petrovic et al., 2016). To investigate whether this is also the case for *Drosophila*, I performed similar experiments on DmMis12C (see Figure 33). The profile of headless DmMis12C shifted to the left when it was incubated with CENP-C (see Figure 33A), similarly to the native DmMis12C (see Figure 24A). As in humans, Head 2 alone does not bind directly to CENP-C (see Figure 33B). I conclude that, as in humans, Head 2 in DmMis12C is not essential for binding with CENP-C. My further SEC experiments also showed there was no tight binding between the Head 2 and the other parts of the DmMis12C (see Figure 35B).

Our studies in humans and other studies in yeast indicate that Dsn1 is the substrate of Aurora B and that the function of Head 2 is in regulating the binding of Mis12C and CENP-C (Dimitrova et al., 2016; Petrovic et al., 2016). The former section indicated that Head 2 may be structurally conserved in *Drosophila*. Thus, it will be interesting to test whether the regulatory function of Head2 is conserved in *Drosophila* or not. I am currently performing tests to determine whether Head 2 in *Drosophila* is also a substrate of Aurora B. Sequence alignment of *Drosophila* Head 2 with the Dsn1 fragments containing the Aurora B substrates from different species indicates sequence similarities, especially the sequence around Ser100 and Ser109 in humans (see Figure 34 around the red arrows). *In vitro* phosphorylation experiments demonstrated that Kmn1 is a substrate

of Aurora B (see Figure 35A-B). MS experiments identified phosphorylation on *Drosophila* Head2 residues Ser17, Ser27 and Ser37 (the green stars in Figure 34). These results show that the Head 2 in *Drosophila* Mis12C is the substrate of Aurora B. We therefore speculate that Aurora B regulates the binding affinity of Mis12C to CENP-C in *Drosophila*, like in humans and yeast. For this, we are currently testing apparent binding affinities of the Mis12C for CENP-C before and after phosphorylation of Mis12C by Aurora B. This section has presented some primary functional tests of the N-terminal of Kmn1 (Head 2). Sequence alignment and biochemical tests indicate that, like human Dsn1, the N-terminal region of Kmn1 is the substrate of Aurora B *in vitro*. These results seem to imply that the lost subunit in DmMis12C is not Dsn1, but could rather be Nsl1. Bearing in mind that the *Drosophila* kinetochore protein sequences evolve rapidly (Meraldi et al., 2006c), we then used a remote homology server, HHpred (Soding et al., 2005), to search homologs of Kmn1 from the family of Drosophilids. Dsn1, rather than Nsl1, was always found as the counterpart in humans, which confirmed our hypothesis.

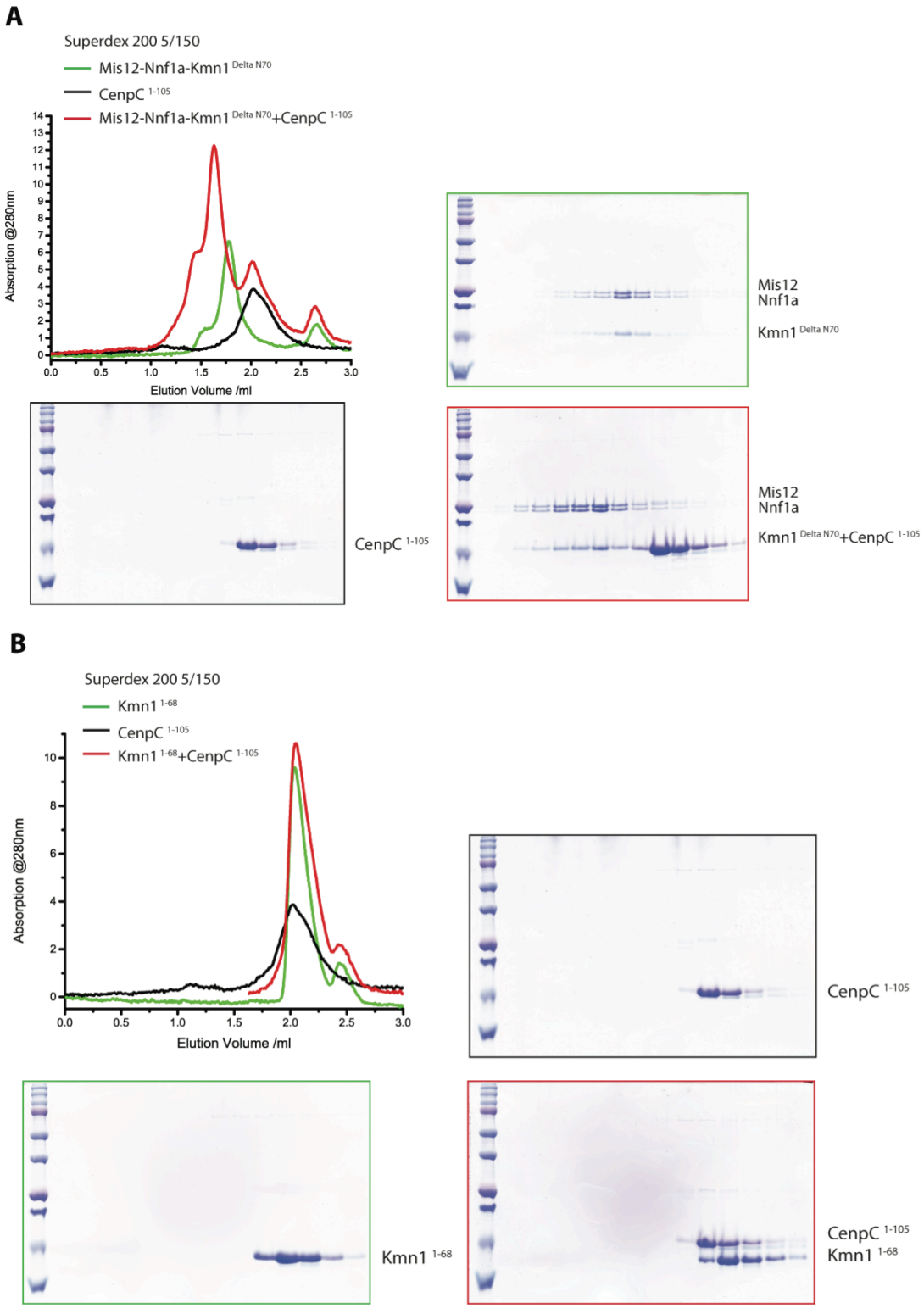

Figure 33 SEC analyses of the indicated DmMis12C and The N-terminal of CENP-C

The fractions were analyzed using SDS-PAGE. (**A**) The shift to the left at the profile of Mis12-Nnf1a-Kmn1$^{\text{Delta N70}}$ indicates CENP-C$^{1\text{-}105}$ binds to this section; (**B**) There is no shift for the profiles of both CENP-C$^{1\text{-}105}$ and Kmn1$^{1\text{-}68}$, which implies no strong binding between these two sections.

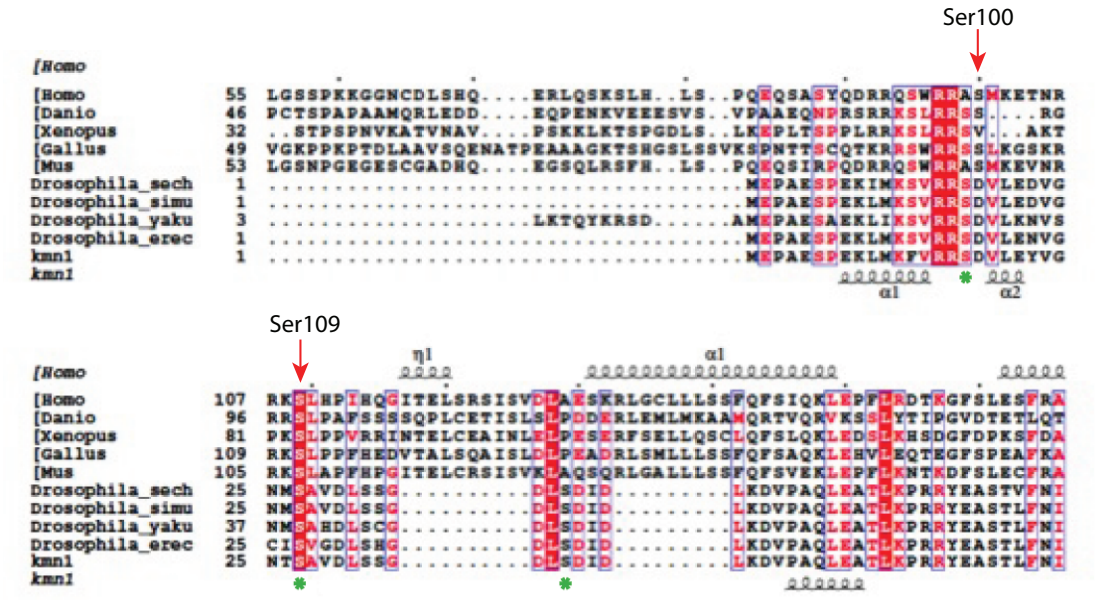

Figure 34 Sequence alignment of *Drosophila* Head2 with the fragment of Dsn1 that contains the Aurora B substrate

The red arrow indicates Ser100 and Ser109, which are the phosphorylation substrates of Aurora B. The green stars denote three phosphorylated Serines (Ser17, Ser27, Ser37) by Aurora B, as identified by MS.

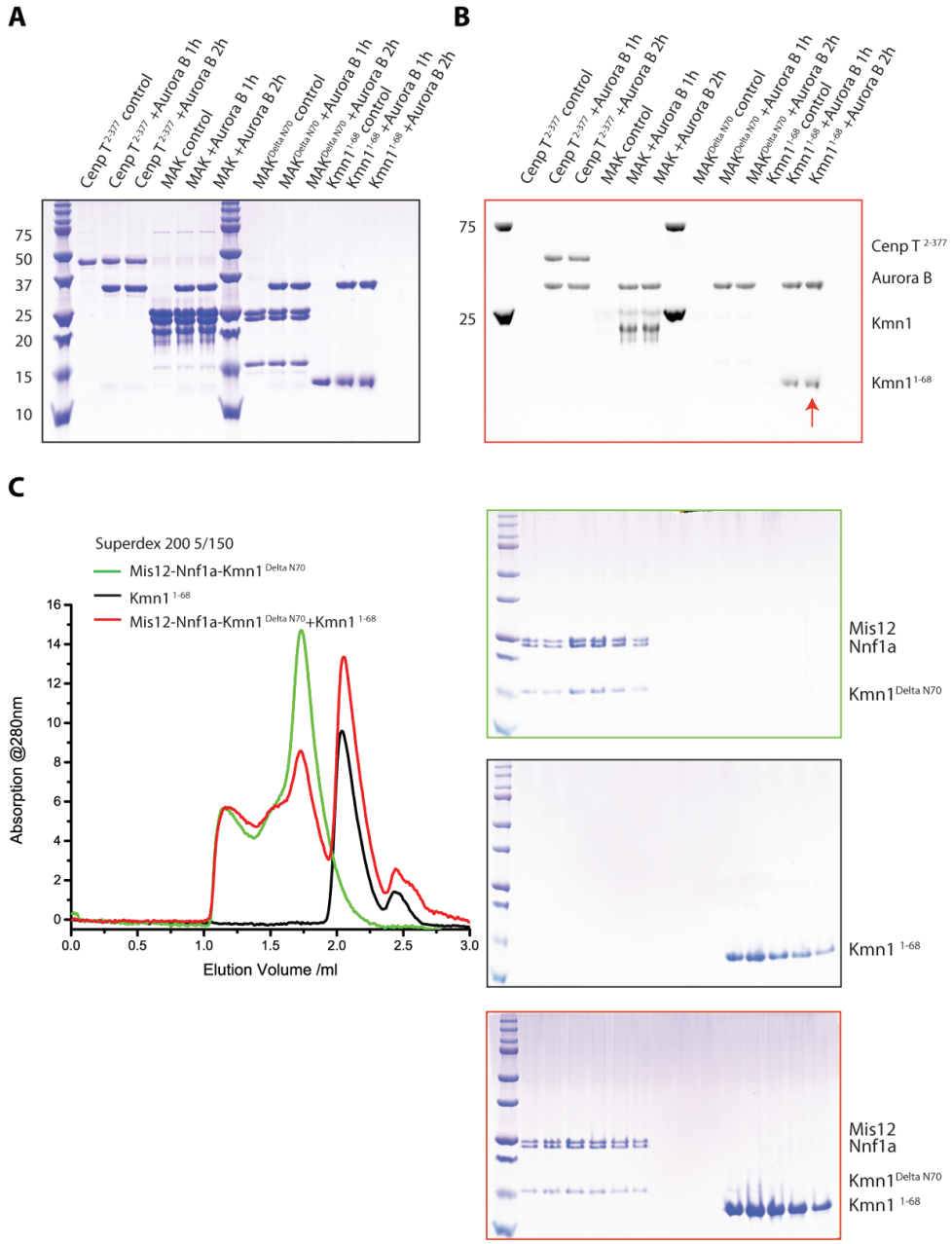

Figure 35 Phosphorylation and additional binding assays of Head2 and headless DmMis12C

(**A**) SDS-PAGE analysis of the samples that were used for the phosphorylation test. The human CENP-T$^{2-377}$, which can be phosphorylated by Aurora B, is used as a positive control. Samples were incubated with Aurora B for one or two hours; (**B**) Pro-Q Diamond phosphorylation stain of all the samples. Compared with the Coomassie Brilliant Blue stain, the red arrow indicates that Head 2 in Kmn1 is the substrate of Aurora B; (**C**) The profiles from both Head 2 and Headless DmMis12C did not shift to the left, indicating that there is no strong interaction between Head 2 and the other sections of DmMis12C.

4 CONCLUSIONS

Here, I have reported my work on the biochemical reconstitution and biophysical characterization of the *Drosophila* outer kinetochore. My studies were motivated by previous observations that most CCAN subunits, with the exception of CENP-C, Zwint (subunit of Knl1/Spc105R complex), and one subunit of Mis12 complex, may not be present in *Drosophila*, (reviewed in (Drinnenberg et al., 2016)). It was therefore of interest to understand how such a simplified *Drosophila* kinetochore retained the core functions of the more complicated kinetochore assemblies in human and yeast. Furthermore, by reconstituting this simpler kinetochore, we hoped to gain access to new sample for structural analysis.

Most of this work has been published in Liu et al., 2016 and Petrovic et al., 2016. Additional work, like the description of the function of Head2 in *Drosophila*, is also included in this thesis but remains unpublished. From the results presented in the previous chapter, I can conclude the following:

1. The overall structure of Mis12C is conserved in *Drosophila*, yeast and humans. However, our biochemical results indicate that unlike humans and yeast, *Drosophila* has two highly similar Mis12 complexes, both of which contain only three subunits. Despite this difference in the number of subunits in the DmMis12 complex and the fact that the primary amino acids sequences differ significantly, the overall structural features of the DmMis12 complex have been conserved from *Drosophila* to yeast and humans. We surmise that this structural conservation underlies a requirement for force sensing and transmission during the process of chromosome segregation, but this hypothesis will require thorough investigations. The loss of one subunit in the Drosophila complex may be connected to the loss of the CCAN components

and the fact that this does not appear to affect the overall stability or fold of the Mis12 complex. Otherwise one would have expected it to be retained.

2. For the DmMis12C, the binding with both CENP-C and Knl1 is also conserved. The two Mis12 complexes in *Drosophila*, the DmMis12a and the DmMis12b complexes, interact tightly with the N-terminal region of CENP-C and also with the C-terminal region of Knl1/Spc105R, suggesting that they have similar interaction potentials. None of the analyses included in our research revealed significant differences in the behavior of the DmMis12a and the DmMis12b complexes. However, previous studies have demonstrated different developmental expression patterns for Nnf1a and Nnf1b (Przewloka et al., 2007; Schittenhelm et al., 2007), raising the possibility of a functional specialization of the two complexes. The C-terminal region of CENP-C directly interacts with CENP-A (Carroll et al., 2010b; Kato et al., 2013; Przewloka et al., 2011b), whereas the N-terminal region connects the inner kinetochore to the Mis12 outer kinetochore complex. There likely is high selective pressure to maintain these interactions through evolution. However, in the vertebrates, CENP-T also contributes to the localization of the Mis12 complex to the kinetochore and provides a second mechanism to link the inner and the outer kinetochore (Huis In 't Veld et al., 2016; Rago et al., 2015a; Suzuki et al., 2015). In *Drosophila*, the CENP-C dependent kinetochore assembly pathway seems to be solely responsible for the recruitment of the outer kinetochore, since there are no other CCAN subunits.

3. Contrary to a previous hypothesis, Kmn1 appears to be the ortholog of Dsn1, not Nsl1. Thus Nsl1, not Dsn1, may be the missing subunit in the *Drosophila* Mis12C. Our contention rests on the observation that the sequence of the N-terminal region of Kmn1, associated with Head1 and with an Aurora B-

regulated region, appears to be structurally conserved. As in human Dsn1, the Head 2 region is the substrate of Aurora B. This strongly suggests that the Kmn1 in DmMis12C functionally resembles Dsn1 in the human Mis12C. Consistently, in *Bombyx mori*, which also lost most of CCAN subunits, Dsn1 is present, whereas Nsl1 hasn't been found (Drinnenberg et al., 2014; Drinnenberg et al., 2016; Mon et al., 2017). This implies that the functions of Dsn1 are much more conserved during evolution.

Ultimately, the overall conclusion emerging from these studies is that *Drosophila* contains highly simplified kinetochores, which display considerable evolutionary and structural plasticity. This simplified complete KMN network could provide a basis for understanding the core functions of the outer kinetochore that are conserved during evolution. However, there are still many open questions. The functional differences of two paralogues of Nnf1 are still not clear. Furthermore, the binding details of both Ndc80C and Knl1 to Mis12C remain unexplained. Further studies will address these questions.

5 REFERENCES

Akiyoshi, B., Nelson, C.R., and Biggins, S. (2013). The Aurora B Kinase Promotes Inner and Outer Kinetochore Interactions in Budding Yeast. Genetics *194*, 785-+.

Albert, B., Bray, D., Hopkin, K., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2014). Essential Cell Biology, Fourth Edition edn (Garland Science).

Alberts, B., Johnson, A., Lewis, J., Morgan, D., Raff, M., Roberts, K., and Walter, P. (2014). Molecular Biology of the Cell, 6th Edition edn (New York and Abindon , UK: Garland Science).

Aldrup-Macdonald, M.E., and Sullivan, B.A. (2014). The past, present, and future of human centromere genomics. Genes *5*, 33-50.

Allshire, R.C., and Karpen, G.H. (2008). Epigenetic regulation of centromeric chromatin: old dogs, new tricks? Nature reviews. Genetics *9*, 923-937.

Altunkaya, G.P., Malvezzi, F., Demianova, Z., Zimniak, T., Litos, G., Weissmann, F., Mechtler, K., Herzog, F., and Westermann, S. (2016). CCAN Assembly Configures Composite Binding Interfaces to Promote Cross-Linking of Ndc80 Complexes at the Kinetochore. Current Biology *26*, 2370-2378.

Alushin, G.M., Musinipally, V., Matson, D., Tooley, J., Stukenberg, P.T., and Nogales, E. (2012). Multimodal microtubule binding by the Ndc80 kinetochore complex. Nat Struct Mol Biol *19*, 1161-+.

Alushin, G.M., Ramey, V.H., Pasqualato, S., Ball, D.A., Grigorieff, N., Musacchio, A., and Nogales, E. (2010). The Ndc80 kinetochore complex forms oligomeric arrays along microtubules. Nature *467*, 805-U868.

Amano, M., Suzuki, A., Hori, T., Backer, C., Okawa, K., Cheeseman, I.M., and Fukagawa, T. (2009). The CENP-S complex is essential for the stable assembly of outer kinetochore structure. The Journal of cell biology *186*, 173-182.

Ando, S., Yang, H., Nozaki, N., Okazaki, T., and Yoda, K. (2002). CENP-A,-B, and-C chromatin complex that contains the I-type α-satellite array constitutes the prekinetochore in HeLa cells. Molecular and Cellular Biology *22*, 2229-2241.

Asbury, C.L., Akiyoshi, B., Sarangapani, K.K., Powers, A.F., Nelson, C.R., Reichow, S.L., Gonen, T., Ranish, J.A., and Biggins, S. (2011). Tension Directly Stabilizes Reconstituted Kinetochore-Microtubule Attachments. Biophysical journal *100*, 531a.

Barth, T.K., Schade, G.O., Schmidt, A., Vetter, I., Wirth, M., Heun, P., Thomae, A.W., and Imhof, A. (2014). Identification of novel Drosophila centromere-associated proteins. Proteomics *14*, 2167-2178.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical biochemistry *72*, 248-254.

Brinkley, B.R., and Stubblefield, E. (1966). The fine structure of the kinetochore of a mammalian cell in vitro. Chromosoma *19*, 28-43.

Cardinale, S., Bergmann, J.H., Kelly, D., Nakano, M., Valdivia, M.M., Kimura, H., Masumoto, H., Larionov, V., and Earnshaw, W.C. (2009). Hierarchical Inactivation of a Synthetic Human Kinetochore by a Chromatin Modifier. Molecular Biology of the Cell *20*, 4194-4204.

Carroll, C.W., Milks, K.J., and Straight, A.F. (2010a). Dual recognition of CENP-A nucleosomes is required for centromere assembly. Journal of Cell Biology *189*, 1143-1155.

Carroll, C.W., Milks, K.J., and Straight, A.F. (2010b). Dual recognition of CENP-A nucleosomes is required for centromere assembly. The Journal of cell biology, jcb-201001013.

Cheeseman, I.M., Chappie, J.S., Wilson-Kubalek, E.M., and Desai, A. (2006). The conserved KMN network constitutes the core microtubule-binding site of the kinetochore. Cell *127*, 983-997.

Cheeseman, I.M., and Desai, A. (2008a). Molecular architecture of the kinetochore-microtubule interface. Nature reviews. Molecular cell biology *9*, 33-46.

Cheeseman, I.M., and Desai, A. (2008b). Molecular architecture of the kinetochore–microtubule interface. Nature reviews Molecular cell biology *9*, 33-46.

Cheeseman, I.M., Hori, T., Fukagawa, T., and Desai, A. (2008). KNL1 and the CENP-H/I/K complex coordinately direct kinetochore assembly in vertebrates. Molecular Biology of the Cell *19*, 587-594.

Cheeseman, I.M., Niessen, S., Anderson, S., Hyndman, F., Yates, J.R., Oegema, K., and Desai, A. (2004). A conserved protein network controls assembly of the outer kinetochore and its ability to sustain tension. Genes & development *18*, 2255-2268.

Ciferri, C., De Luca, J., Monzani, S., Ferrari, K.J., Ristic, D., Wyman, C., Stark, H., Kilmartin, J., Salmon, E.D., and Musacchio, A. (2005). Architecture of the human Ndc80-Hec1 complex, a critical constituent of the outer kinetochore. Journal of Biological Chemistry *280*, 29088-29095.

Ciferri, C., Pasqualato, S., Screpanti, E., Varetti, G., Santaguida, S., Dos Reis, G., Maiolica, A., Polka, J., De Luca, J.G., De Wulf, P.*, et al.* (2008). Implications for kinetochore-microtubule attachment from the structure of an engineered Ndc80 complex. Cell *133*, 427-439.

Cleveland, D.W., Mao, Y., and Sullivan, K.F. (2003a). Centromeres and kinetochores: from epigenetics to mitotic checkpoint signaling. Cell *112*, 407-421.

Cleveland, D.W., Mao, Y.H., and Sullivan, K.F. (2003b). Centromeres and kinetochores: From epigenetics to mitotic checkpoint signaling. Cell *112*, 407-421.

Cohen, R., Espelin, C., De Wulf, P., Sorger, P., Harrison, S., and Simons, K. (2008). Structural and functional dissection of Mif2p, a conserved DNA-binding kinetochore protein. Molecular biology of the cell *19*, 4480-4491.

De Rop, V., Padeganeh, A., and Maddox, P.S. (2012). CENP-A: the key player behind centromere identity, propagation, and kinetochore assembly. Chromosoma *121*, 527-538.

DeLuca, J.G., Gall, W.E., Ciferri, C., Cimini, D., Musacchio, A., and Salmon, E.D. (2006). Kinetochore microtubule dynamics and attachment stability are regulated by Hec1. Cell *127*, 969-982.

DeLuca, K.F., Lens, S.M.A., and DeLuca, J.G. (2011). Temporal changes in Hec1 phosphorylation control kinetochore-microtubule attachment stability during mitosis. J Cell Sci *124*, 622-634.

Desai, A., Verma, S., Mitchison, T.J., and Walczak, C.E. (1999). Kin I kinesins are microtubule-destabilizing enzymes. Cell *96*, 69-78.

Dimitrova, Y.N., Jenni, S., Valverde, R., Khin, Y., and Harrison, S.C. (2016). Structure of the MIND Complex Defines a Regulatory Focus for Yeast Kinetochore Assembly. Cell *167*, 1014-+.

Drinnenberg, I.A., deYoung, D., Henikoff, S., and Malik, H.S. (2014). Recurrent loss of CenH3 is associated with independent transitions to holocentricity in insects. Elife *3*.

Drinnenberg, I.A., Henikoff, S., and Malik, H.S. (2016). Evolutionary Turnover of Kinetochore Proteins: A Ship of Theseus? Trends in Cell Biology *26*, 498-510.

Drozdetskiy, A., Cole, C., Procter, J., and Barton, G.J. (2015). JPred4: a protein secondary structure prediction server. Nucleic acids research *43*, W389-W394.

Dutcher, S.K. (2001). The tubulin fraternity: alpha to eta. Curr Opin Cell Biol *13*, 49-54.

Earnshaw, W.C. (2015). Discovering centromere proteins: from cold white hands to the A, B, C of CENPs. Nature reviews. Molecular cell biology *16*, 443-449.

Earnshaw, W.C., and Rothfield, N. (1985). Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. Chromosoma *91*, 313-321.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res *32*, 1792-1797.

Erhardt, S., Mellone, B.G., Betts, C.M., Zhang, W.G., Karpen, G.H., and Straight, A.F. (2008). Genome-wide analysis reveals a cell cycle-dependent mechanism controlling centromere propagation. Journal of Cell Biology *183*, 805-818.

Espeut, J., Cheerambathur, D.K., Krenning, L., Oegema, K., and Desai, A. (2012). Microtubule binding by KNL-1 contributes to spindle checkpoint silencing at the kinetochore. Journal of Cell Biology *196*, 469-482.

Foley, E.A., and Kapoor, T.M. (2013). Microtubule attachment and spindle assembly checkpoint signalling at the kinetochore. Nature Reviews Molecular Cell Biology *14*, 25-37.

Foltz, D.R., Jansen, L.E., Black, B.E., Bailey, A.O., Yates, J.R., and Cleveland, D.W. (2006). The human CENP-A centromeric nucleosome-associated complex. Nature cell biology *8*, 458-469.

Fukagawa, T., and Brown, W.R.A. (1997). Efficient conditional mutation of the vertebrate CENP-C gene. Human molecular genetics *6*, 2301-2308.

Funabiki, H., and Wynne, D.J. (2013). Making an effective switch at the kinetochore by phosphorylation and dephosphorylation. Chromosoma *122*, 135-158.

Garbett, N.C., Mekmaysy, C.S., and Chaires, J.B. (2010). Sedimentation Velocity Ultracentrifugation Analysis for Hydrodynamic Characterization of G-Quadruplex Structures. G-Quadruplex DNA: Methods and Protocols *608*, 97-120.

Gascoigne, K.E., and Cheeseman, I.M. (2011). Kinetochore assembly: if you build it, they will come. Curr Opin Cell Biol *23*, 102-108.

Gassmann, R., Holland, A.J., Varma, D., Wan, X.H., Civril, F., Cleveland, D.W., Oegema, K., Salmon, E.D., and Desai, A. (2010). Removal of Spindly from microtubule-attached kinetochores controls spindle checkpoint silencing in human cells. Genes & Development *24*, 957-971.

Gould, G.W. (2016). Animal cell cytokinesis: The role of dynamic changes in the plasma membrane proteome and lipidome. Semin Cell Dev Biol *53*, 64-73.

Grimm, M., Zimniak, T., Kahraman, A., and Herzog, F. (2015). xVis: a web server for the schematic visualization and interpretation of crosslink-derived spatial restraints. Nucleic acids research *43*, W362-369.

Guimaraes, G.J., Dong, Y.M., McEwen, B.F., and DeLuca, J.G. (2008). Kinetochore-Microtubule Attachment Relies on the Disordered N-Terminal Tail Domain of Hec1. Current Biology *18*, 1778-1784.

Guse, A., Carroll, C.W., Moree, B., Fuller, C.J., and Straight, A.F. (2011). In vitro centromere and kinetochore assembly on defined chromatin templates. Nature *477*, 354-358.

Heeger, S., Leismann, O., Schittenhelm, R., Schraidt, O., Heidmann, S., and Lehner, C.F. (2005a). Genetic interactions of separase regulatory subunits reveal the diverged Drosophila Cenp-C homolog. Genes & Development *19*, 2041-2053.

Heeger, S., Leismann, O., Schittenhelm, R., Schraidt, O., Heidmann, S., and Lehner, C.F. (2005b). Genetic interactions of separase regulatory subunits reveal the diverged Drosophila Cenp-C homolog. Genes Dev *19*, 2041-2053.

Herzog, F., Kahraman, A., Boehringer, D., Mak, R., Bracher, A., Walzthoeni, T., Leitner, A., Beck, M., Hartl, F.U., Ban, N.*, et al.* (2012). Structural probing of a protein phosphatase 2A network by chemical cross-linking and mass spectrometry. Science *337*, 1348-1352.

Holland, S., Ioannou, D., Haines, S., and Brown, W.R. (2005). Comparison of Dam tagging and chromatin immunoprecipitation as tools for the identification of the binding sites for S. pombe CENP-C. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology *13*, 73-83.

Holt, L.J., Tuch, B.B., Villen, J., Johnson, A.D., Gygi, S.P., and Morgan, D.O. (2009). Global Analysis of Cdk1 Substrate Phosphorylation Sites Provides Insights into Evolution. Science *325*, 1682-1686.

Hori, T., Okada, M., Maenaka, K., and Fukagawa, T. (2008). CENP-O class proteins form a stable complex and are required for proper kinetochore function. Molecular biology of the cell *19*, 843-854.

Hori, T., Shang, W.-H., Takeuchi, K., and Fukagawa, T. (2013). The CCAN recruits CENP-A to the centromere and forms the structural core for kinetochore assembly. J Cell Biol *200*, 45-60.

Hornung, P., Maier, M., Alushin, G.M., Lander, G.C., Nogales, E., and Westermann, S. (2011a). Molecular architecture and connectivity of the budding yeast Mtw1 kinetochore complex. Journal of molecular biology *405*, 548-559.

Hornung, P., Maier, M., Alushin, G.M., Lander, G.C., Nogales, E., and Westermann, S. (2011b). Molecular architecture and connectivity of the budding yeast Mtw1 kinetochore complex. J Mol Biol *405*, 548-559.

Hornung, P., Troc, P., Malvezzi, F., Maier, M., Demianova, Z., Zimniak, T., Litos, G., Lampert, F., Schleiffer, A., Brunner, M., *et al.* (2014a). A cooperative mechanism drives budding yeast kinetochore assembly downstream of CENP-A. The Journal of cell biology *206*, 509-524.

Hornung, P., Troc, P., Malvezzi, F., Maier, M., Demianova, Z., Zimniak, T., Litos, G., Lampert, F., Schleiffer, A., Brunner, M., *et al.* (2014b). A cooperative mechanism drives budding yeast kinetochore assembly downstream of CENP-A. Journal of Cell Biology *206*, 509-524.

Howard, J., and Hyman, A.A. (2003). Dynamics and mechanics of the microtubule plus end. Nature *422*, 753-758.

Howlett, G.J., Minton, A.P., and Rivas, G. (2006). Analytical ultracentrifugation for association and assembly the study of protein. Curr Opin Chem Biol *10*, 430-436.

Huis In 't Veld, P.J., Jeganathan, S., Petrovic, A., Singh, P., John, J., Krenn, V., Weissmann, F., Bange, T., and Musacchio, A. (2016). Molecular basis of outer kinetochore assembly on CENP-T. Elife *5*.

Izuta, H., Ikeno, M., Suzuki, N., Tomonaga, T., Nozaki, N., Obuse, C., Kisu, Y., Goshima, N., Nomura, F., and Nomura, N. (2006). Comprehensive analysis of the ICEN (interphase centromere complex) components enriched in the CENP−A chromatin of human cells. Genes to Cells *11*, 673-684.

Joglekar, A.P., Bloom, K., and Salmon, E.D. (2009). In vivo protein architecture of the eukaryotic kinetochore with nanometer scale accuracy. Current Biology *19*, 694-699.

Karsenti, E., and Vernos, I. (2001). The mitotic spindle: a self-made machine. Science *294*, 543-547.

Kato, H., Jiang, J., Zhou, B.-R., Rozendaal, M., Feng, H., Ghirlando, R., Xiao, T.S., Straight, A.F., and Bai, Y. (2013). A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. Science *340*, 1110-1113.

Kim, S., and Yu, H.T. (2015). Multiple assembly mechanisms anchor the KMN spindle checkpoint platform at human mitotic kinetochores. Journal of Cell Biology *208*, 181-196.

Kitagawa, K., and Hieter, P. (2001). Evolutionary conservation between budding yeast and human kinetochores. Nature Reviews Molecular Cell Biology *2*, 678-687.

Kiyomitsu, T., Iwasaki, O., Obuse, C., and Yanagida, M. (2010). Inner centromere formation requires hMis14, a trident kinetochore protein that specifically recruits HP1 to human chromosomes. The Journal of cell biology *188*, 791-807.

Kiyomitsu, T., Murakami, H., and Yanagida, M. (2011). Protein Interaction Domain Mapping of Human Kinetochore Protein Blinkin Reveals a Consensus Motif for Binding of Spindle Assembly Checkpoint Proteins Bub1 and BubR1. Molecular and Cellular Biology *31*, 998-1011.

Kiyomitsu, T., Obuse, C., and Yanagida, M. (2007). Human Blinkin/AF15q14 is required for chromosome alignment and the mitotic checkpoint through direct interaction with Bub1 and BubR1. Developmental Cell *13*, 663-676.

Klare, K., Weir, J.R., Basilico, F., Zimniak, T., Massimiliano, L., Ludwigs, N., Herzog, F., and Musacchio, A. (2015a). CENP-C is a blueprint for constitutive centromere-associated network assembly within human kinetochores. The Journal of cell biology *210*, 11-22.

Klare, K., Weir, J.R., Basilico, F., Zimniak, T., Massimiliano, L., Ludwigs, N., Herzog, F., and Musacchio, A. (2015b). CENP-C is a blueprint for constitutive centromere–associated network assembly within human kinetochores. J Cell Biol, jcb. 201412028.

Kline, S.L., Cheeseman, I.M., Hori, T., Fukagawa, T., and Desai, A. (2006). The human Mis12 complex is required for kinetochore assembly and proper chromosome segregation. The Journal of cell biology *173*, 9-17.

Kops, G.J.P.L., Kim, Y., Weaver, B.A.A., Mao, Y.H., McLeod, I., Yates, J.R., Tagaya, M., and Cleveland, D.W. (2005). ZW10 links mitotic checkpoint signaling to the structural kinetochore. Journal of Cell Biology *169*, 49-60.

Krenn, V., Wehenkel, A., Li, X.Z., Santaguida, S., and Musacchio, A. (2012). Structural analysis reveals features of the spindle checkpoint kinase Bub1-kinetochore subunit Knl1 interaction. Journal of Cell Biology *196*, 451-467.

Kwon, M.S., Hori, T., Okada, M., and Fukagawa, T. (2007). CENP-C is involved in chromosome segregation, mitotic checkpoint function, and kinetochore assembly. Molecular Biology of the Cell *18*, 2155-2168.

Lara-Gonzalez, P., Westhorpe, F.G., and Taylor, S.S. (2012a). The spindle assembly checkpoint. Current biology : CB *22*, R966-980.

Lara-Gonzalez, P., Westhorpe, F.G., and Taylor, S.S. (2012b). The Spindle Assembly Checkpoint. Current Biology *22*, R966-R980.

Laue, T.M. (1995). Sedimentation equilibrium as thermodynamic tool. Method Enzymol *259*, 427-452.

Liu, D., Vleugel, M., Backer, C.B., Hori, T., Fukagawa, T., Cheeseman, I.M., and Lampson, M.A. (2010). Regulated targeting of protein phosphatase 1 to the outer kinetochore by KNL1 opposes Aurora B kinase. Journal of Cell Biology *188*, 809-820.

Liu, S.T., Rattner, J.B., Jablonski, S.A., and Yen, T.J. (2006). Mapping the assembly pathways that specify formation of the trilaminar kinetochore plates in human cells. Journal of Cell Biology *175*, 41-53.

Liu, Y., Petrovic, A., Rombaut, P., Mosalaganti, S., Keller, J., Raunser, S., Herzog, F., and Musacchio, A. (2016). Insights from the reconstitution of the divergent outer kinetochore of Drosophila melanogaster. Open biology *6*.

Lodish, H., Berk, A., Kaiser, C.A., Krieger, M., Scott, M.P., Bretscher, A., Ploegh, H., and Matsudaira, P. (2007). Molecular Cell Biology, 6th Edition edn (W. H. Freeman and Company).

Lupas, A., Van Dyke, M., and Stock, J. (1991). Predicting coiled coils from protein sequences. Science *252*, 1162-1164.

Malvezzi, F., Litos, G., Schleiffer, A., Heuck, A., Mechtler, K., Clausen, T., and Westermann, S. (2013). A structural basis for kinetochore recruitment of the Ndc80 complex via two distinct centromere receptors. Embo J *32*, 409-423.

Marshall, O.J., Chueh, A.C., Wong, L.H., and Choo, K.H.A. (2008). Neocentromeres: New Insights into Centromere Structure, Disease Development, and Karyotype Evolution. The American Journal of Human Genetics *82*, 261-282.

Maskell, D.P., Hu, X.-W., and Singleton, M.R. (2010a). Molecular architecture and assembly of the yeast kinetochore MIND complex. The Journal of cell biology *190*, 823-834.

Maskell, D.P., Hu, X.W., and Singleton, M.R. (2010b). Molecular architecture and assembly of the yeast kinetochore MIND complex. The Journal of cell biology *190*, 823-834.

Meraldi, P., McAinsh, A.D., Rheinbay, E., and Sorger, P.K. (2006a). Phylogenetic and structural analysis of centromeric DNA and kinetochore proteins. Genome biology *7*, R23.

Meraldi, P., McAinsh, A.D., Rheinbay, E., and Sorger, P.K. (2006b). Phylogenetic and structural analysis of centromeric DNA and kinetochore proteins. Genome Biol *7*, R23.

Meraldi, P., McAinsh, A.D., Rheinbay, E., and Sorger, P.K. (2006c). Phylogenetic and structural analysis of centromeric DNA and kinetochore proteins. Genome Biology *7*.

Milks, K.J., Moree, B., and Straight, A.F. (2009). Dissection of CENP-C-directed Centromere and Kinetochore Assembly. Molecular Biology of the Cell *20*, 4246-4255.

Mimori-Kiyosue, Y., Grigoriev, I., Lansbergen, G., Sasaki, H., Matsui, C., Severin, F., Galjart, N., Grosveld, F., Vorobjev, I., Tsukita, S.*, et al.* (2005). CLASP1 and CLASP2 bind to EB1 and regulate microtubule plus-end dynamics at the cell cortex. Journal of Cell Biology *168*, 141-153.

Mitchison, T., Evans, L., Schulze, E., and Kirschner, M. (1986). Sites of Microtubule Assembly and Disassembly in the Mitotic Spindle. Cell *45*, 515-527.

Mon, H., Lee, J.M., Sato, M., and Kusakabe, T. (2017). Identification and functional analysis of outer kinetochore genes in the holocentric insect Bombyx mori. Insect biochemistry and molecular biology *86*, 1-8.

Moore, L.L., and Roth, M.B. (2001). HCP-4, a CENP-C-like protein in Caenorhabditis elegans, is required for resolution of sister centromeres. The Journal of cell biology *153*, 1199-1208.

Musacchio, A., and Desai, A. (2017). A Molecular View of Kinetochore Assembly and Function. Biology *6*.

Musacchio, A., and Salmon, E.D. (2007). The spindle-assembly checkpoint in space and time. Nature Reviews Molecular Cell Biology *8*, 379-393.

Nagpal, H., Hori, T., Furukawa, A., Sugase, K., Osakabe, A., Kurumizaka, H., and Fukagawa, T. (2015a). Dynamic changes in CCAN organization through CENP-C during cell-cycle progression. Molecular biology of the cell *26*, 3768-3776.

Nagpal, H., Hori, T., Furukawa, A., Sugase, K., Osakabe, A., Kurumizaka, H., and Fukagawa, T. (2015b). Dynamic changes in the CCAN organization through CENP-C during cell-cycle progression. Molecular biology of the cell.

Nakano, M., Cardinale, S., Noskov, V.N., Gassmann, R., Vagnarelli, P., Kandels-Lewis, S., Larionov, V., Earnshaw, W.C., and Masumoto, H. (2008). Inactivation of a human kinetochore by specific targeting of chromatin modifiers. Dev Cell *14*, 507-522.

Nekrasov, V.S., Smith, M.A., Peak-Chew, S., and Kilmartin, J.V. (2003). Interactions between centromere complexes in Saccharomyces cerevisiae. Molecular Biology of the Cell *14*, 4931-4946.

Nishino, T., Rago, F., Hori, T., Tomii, K., Cheeseman, I.M., and Fukagawa, T. (2013). CENP-T provides a structural platform for outer kinetochore assembly. Embo J *32*, 424-436.

Nogales, E., and Wang, H.W. (2006). Structural intermediates in microtubule assembly and disassembly: how and why? Curr Opin Cell Biol *18*, 179-184.

Norbury, C., and Nurse, P. (1992). Animal-Cell Cycles and Their Control. Annual review of biochemistry *61*, 441-470.

Obuse, C., Iwasaki, O., Kiyomitsu, T., and Yanagida, M. (2005). A conserved Mis12 centromere complex is linked to heterochromatic HP1 and outer kinetochore proteins Zwint-1 and HEC1. Cell Struct Funct *30*, 7-7.

Oegema, K., Desai, A., Rybina, S., Kirkham, M., and Hyman, A.A. (2001a). Functional analysis of kinetochore assembly in Caenorhabditis elegans. The Journal of cell biology *153*, 1209-1226.

Oegema, K.F., Desai, A., Rybina, S., and Hyman, A. (2001b). Functional analysis of kinetochore assembly in C. elegans. Molecular Biology of the Cell *12*, 177A-177A.

Okada, M., Cheeseman, I.M., Hori, T., Okawa, K., McLeod, I.X., Yates, J.R., Desai, A., and Fukagawa, T. (2006). The CENP-H–I complex is required for the efficient incorporation of newly synthesized CENP-A into centromeres. Nature cell biology *8*, 446-457.

Orr, B., and Sunkel, C.E. (2011). Drosophila CENP-C is essential for centromere identity. Chromosoma *120*, 83-96.

Pagliuca, C., Draviam, V.M., Marco, E., Sorger, P.K., and De Wulf, P. (2009). Roles for the Conserved Spc105p/Kre28p Complex in Kinetochore-Microtubule Binding and the Spindle Assembly Checkpoint. PloS one *4*.

Panchenko, T., Sorensen, T.C., Woodcock, C.L., Kan, Z.Y., Wood, S., Resch, M.G., Luger, K., Englander, S.W., Hansen, J.C., and Black, B.E. (2011). Replacement of histone H3 with CENP-A directs global nucleosome array condensation and loosening of nucleosome superhelical termini. Proceedings of the National Academy of Sciences of the United States of America *108*, 16588-16593.

Pavletich, N.P. (1999). Mechanisms of cyclin-dependent kinase regulation: Structures of Cdks, their cyclin activators, and Cip and INK4 inhibitors. Journal of Molecular Biology *287*, 821-828.

Petrovic, A., Keller, J., Liu, Y.H., Overlack, K., John, J., Dimitrova, Y.N., Jenni, S., van Gerwen, S., Stege, P., Wohlgemuth, S.*, et al.* (2016). Structure of the MIS12 Complex and Molecular Basis of Its Interaction with CENP-C at Human Kinetochores. Cell *167*, 1028-+.

Petrovic, A., Mosalaganti, S., Keller, J., Mattiuzzo, M., Overlack, K., Krenn, V., De Antoni, A., Wohlgemuth, S., Cecatiello, V., Pasqualato, S.*, et al.* (2014a). Modular assembly of RWD domains on the Mis12 complex underlies outer kinetochore organization. Molecular cell *53*, 591-605.

Petrovic, A., Mosalaganti, S., Keller, J., Mattiuzzo, M., Overlack, K., Krenn, V., De Antoni, A., Wohlgemuth, S., Cecatiello, V., Pasqualato, S.*, et al.* (2014b). Modular assembly of RWD domains on the Mis12 complex underlies outer kinetochore organization. Molecular cell *53*, 591-605.

Petrovic, A., Pasqualato, S., Dube, P., Krenn, V., Santaguida, S., Cittaro, D., Monzani, S., Massimiliano, L., Keller, J., Tarricone, A.*, et al.* (2010a). The MIS12 complex is a protein interaction hub for outer kinetochore assembly. The Journal of cell biology *190*, 835-852.

Petrovic, A., Pasqualato, S., Dube, P., Krenn, V., Santaguida, S., Cittaro, D., Monzani, S., Massimiliano, L., Keller, J., Tarricone, A.*, et al.* (2010b). The MIS12 complex is a protein interaction hub for outer kinetochore assembly. The Journal of cell biology *190*, 835-852.

Pluta, A.F., Mackay, A.M., Ainsztein, A.M., Goldberg, I.G., and Earnshaw, W.C. (1995). The centromere: hub of chromosomal activities. Science *270*, 1591-1594.

Primorac, I., Weir, J.R., Chiroli, E., Gross, F., Hoffmann, I., van Gerwen, S., Ciliberto, A., and Musacchio, A. (2013). Bub3 reads phosphorylated MELT repeats to promote spindle assembly checkpoint signaling. Elife *2*.

Przewloka, M.R., Venkei, Z., Bolanos-Garcia, V.M., Debski, J., Dadlez, M., and Glover, D.M. (2011a). CENP-C is a structural platform for kinetochore assembly. Current biology : CB *21*, 399-405.

Przewloka, M.R., Venkei, Z., Bolanos-Garcia, V.M., Debski, J., Dadlez, M., and Glover, D.M. (2011b). CENP-C is a structural platform for kinetochore assembly. Current Biology *21*, 399-405.

Przewloka, M.R., Venkei, Z., and Glover, D.M. (2009). Searching for Drosophila Dsn1 kinetochore protein. Cell Cycle *8*, 1292-1293.

Przewloka, M.R., Zhang, W., Costa, P., Archambault, V., D'Avino, P.P., Lilley, K.S., Laue, E.D., McAinsh, A.D., and Glover, D.M. (2007). Molecular analysis of core kinetochore composition and assembly in Drosophila melanogaster. PloS one *2*, e478.

Rago, F., Gascoigne, K.E., and Cheeseman, I.M. (2015a). Distinct organization and regulation of the outer kinetochore KMN network downstream of CENP-C and CENP-T. Current biology : CB *25*, 671-677.

Rago, F., Gascoigne, K.E., and Cheeseman, I.M. (2015b). Distinct Organization and Regulation of the Outer Kinetochore KMN Network Downstream of CENP-C and CENP-T. Current Biology *25*, 671-677.

Richter, M.M., Poznanski, J., Zdziarska, A., Czarnocki-Cieciura, M., Lipinszki, Z., Dadlez, M., Glover, D.M., and Przewloka, M.R. (2016). Network of protein interactions within the Drosophila inner kinetochore. Open biology *6*.

Rieder, C.L. (1982). The Formation, Structure, and Composition of the Mammalian Kinetochore and Kinetochore Fiber. Int Rev Cytol *79*, 1-58.

Rieder, C.L., and Salmon, E.D. (1998). The vertebrate cell kinetochore and its roles during mitosis. Trends in Cell Biology *8*, 310-318.

Sacristan, C., and Kops, G.J.P.L. (2015). Joined at the hip: kinetochores, microtubules, and spindle assembly checkpoint signaling. Trends in Cell Biology *25*, 21-28.

Saitoh, H., Tomkiel, J., Cooke, C.A., Ratrie, H., 3rd, Maurer, M., Rothfield, N.F., and Earnshaw, W.C. (1992). CENP-C, an autoantigen in scleroderma, is a component of the human inner kinetochore plate. Cell *70*, 115-125.

Santaguida, S., and Musacchio, A. (2009). The life and miracles of kinetochores. Embo J *28*, 2511-2531.

Sawin, K.E., Leguellec, K., Philippe, M., and Mitchison, T.J. (1992). Mitotic Spindle Organization by a Plus-End-Directed Microtubule Motor. Nature *359*, 540-543.

Schittenhelm, R.B., Chaleckis, R., and Lehner, C.F. (2009a). Intrakinetochore localization and essential functional domains of Drosophila Spc105. The EMBO Journal *28*, 2374-2386.

Schittenhelm, R.B., Chaleckis, R., and Lehner, C.F. (2009b). Intrakinetochore localization and essential functional domains of Drosophila Spc105. The EMBO journal *28*, 2374-2386.

Schittenhelm, R.B., Heeger, S., Althoff, F., Walter, A., Heidmann, S., Mechtler, K., and Lehner, C.F. (2007). Spatial organization of a ubiquitous eukaryotic kinetochore protein network in Drosophila chromosomes. Chromosoma *116*, 385-402.

Schuck, P. (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. Biophys J *78*, 1606-1619.

Screpanti, E., De Antoni, A., Alushin, G.M., Petrovic, A., Melis, T., Nogales, E., and Musacchio, A. (2011a). Direct Binding of Cenp-C to the Mis12 Complex Joins the Inner and Outer Kinetochore. Current Biology *21*, 391-398.

Screpanti, E., De Antoni, A., Alushin, G.M., Petrovic, A., Melis, T., Nogales, E., and Musacchio, A. (2011b). Direct binding of Cenp-C to the Mis12 complex joins the inner and outer kinetochore. Current biology : CB *21*, 391-398.

Selleck, W., and Tan, S. (2008). Recombinant protein complex expression in E. coli. Current protocols in protein science *Chapter 5*, Unit 5 21.

Soding, J., Biegert, A., and Lupas, A.N. (2005). The HHpred interactive server for protein homology detection and structure prediction. Nucleic acids research *33*, W244-W248.

Song, K., Gronemeyer, B., Lu, W., Eugster, E., and Tomkiel, J.E. (2002). Mutational analysis of the central centromere targeting domain of human centromere protein C, (CENP-C). Exp Cell Res *275*, 81-91.

Srivastava, V., Iglesias, P.A., and Robinson, D.N. (2016). Cytokinesis: Robust cell shape regulation. Semin Cell Dev Biol *53*, 39-44.

Starr, D.A., Saffery, R., Li, Z.X., Simpson, A.E., Choo, K.H.A., Yen, T.J., and Goldberg, M.L. (2000). HZwint-1,a novel human kinetochore component that interacts with HZW10. J Cell Sci *113*, 1939-1950.

Stellfox, M.E., Bailey, A.O., and Foltz, D.R. (2013). Putting CENP-A in its place. Cellular and molecular life sciences : CMLS *70*, 387-406.

Steuer, E.R., Wordeman, L., Schroer, T.A., and Sheetz, M.P. (1990). Localization of Cytoplasmic Dynein to Mitotic Spindles and Kinetochores. Nature *345*, 266-268.

Sugimoto, K., Kuriyama, K., Shibata, A., and Himeno, M. (1997a). Characterization of internal DNA-binding and C-terminal dimerization domains of human centromere/kinetochore autoantigen CENP-C in vitro: role of DNA-binding and self-associating activities in kinetochore organization. Chromosome Research *5*, 132-141.

Sugimoto, K., Kuriyama, K., Shibata, A., and Himeno, M. (1997b). Characterization of internal DNA-binding and C-terminal dimerization domains of human centromere/kinetochore autoantigen CENP-C in vitro: role of DNA-binding and self-associating activities in kinetochore organization. Chromosome Res *5*, 132-141.

Sullivan, B.A., and Karpen, G.H. (2004). Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. Nat Struct Mol Biol *11*, 1076-1083.

Suzuki, A., Badger, B.L., and Salmon, E.D. (2015). A quantitative description of Ndc80 complex linkage to human kinetochores. Nature communications *6*.

Tan, S., Kern, R.C., and Selleck, W. (2005). The pST44 polycistronic expression system for producing protein complexes in Escherichia coli. Protein Expr Purif *40*, 385-395.

Tanaka, K., Chang, H.L., Kagami, A., and Watanabe, Y. (2009). CENP-C Functions as a Scaffold for Effectors with Essential Kinetochore Functions in Mitosis and Meiosis. Developmental Cell *17*, 334-343.

Tomkiel, J., Cooke, C.A., Saitoh, H., Bernat, R.L., and Earnshaw, W.C. (1994). Cenp-C Is Required for Maintaining Proper Kinetochore Size and for a Timely Transition to Anaphase. Journal of Cell Biology *125*, 531-545.

Valverde, R., Ingram, J., and Harrison, S.C. (2016). Conserved Tetramer Junction in the Kinetochore Ndc80 Complex. Cell reports *17*, 1915-1922.

Venkei, Z., Przewloka, M.R., and Glover, D.M. (2011a). Drosophila Mis12 complex acts as a single functional unit essential for anaphase chromosome movement and a robust spindle assembly checkpoint. Genetics *187*, 131-140.

Venkei, Z., Przewloka, M.R., and Glover, D.M. (2011b). Drosophila Mis12 complex acts as a single functional unit essential for anaphase chromosome movement and a robust spindle assembly checkpoint. Genetics *187*, 131-140.

Venkei, Z., Przewloka, M.R., Ladak, Y., Albadri, S., Sossick, A., Juhasz, G., Novak, B., and Glover, D.M. (2012a). Spatiotemporal dynamics of Spc105 regulates the assembly of the Drosophila kinetochore. Open biology *2*, 110032.

Venkei, Z., Przewloka, M.R., Ladak, Y., Albadri, S., Sossick, A., Juhasz, G., Novák, B., and Glover, D.M. (2012b). Spatiotemporal dynamics of Spc105 regulates the assembly of the Drosophila kinetochore. Open biology *2*, 110032.

Vermaak, D., Hayden, H.S., and Henikoff, S. (2002). Centromere targeting element within the histone fold domain of Cid. Molecular and Cellular Biology *22*, 7553-7561.

Walczak, C.E., and Heald, R. (2008). Mechanisms of mitotic spindle assembly and function. International Review of Cytology: A Survey of Cell Biology, Vol 265 *265*, 111-+.

Walzthoeni, T., Claassen, M., Leitner, A., Herzog, F., Bohn, S., Forster, F., Beck, M., and Aebersold, R. (2012). False discovery rate estimation for cross-linked peptides identified by mass spectrometry. Nat Methods *9*, 901-903.

Wang, H., Hu, X., Ding, X., Dou, Z., Yang, Z., Shaw, A.W., Teng, M., Cleveland, D.W., Goldberg, M.L., Niu, L.*, et al.* (2004). Human Zwint-1 specifies localization of Zeste White 10 to kinetochores and is essential for mitotic checkpoint signaling. Journal of Biological Chemistry *279*, 54590-54598.

Warburton, P.E., Cooke, C.A., Bourassa, S., Vafa, O., Sullivan, B.A., Stetten, G., Gimelli, G., Warburton, D., Tyler-Smith, C., Sullivan, K.F.*, et al.* (1997). Immunolocalization of CENP-A suggests a distinct nucleosome structure at the inner kinetochore plate of active centromeres. Current biology : CB *7*, 901-904.

Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M., and Barton, G.J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics *25*, 1189-1191.

Wei, R.R., Al-Bassam, J., and Harrison, S.C. (2007). The Ndc80/HEC1 complex is a contact point for kinetochore-microtubule attachment. Nat Struct Mol Biol *14*, 54-59.

Wei, R.R., Schnell, J.R., Larsen, N.A., Sorger, P.K., Chou, J.J., and Harrison, S.C. (2006). Structure of a central component of the yeast kinetochore: The Spc24p/Spc25p globular domain. Structure *14*, 1003-1009.

Wei, R.R., Sorger, P.K., and Harrison, S.C. (2005). Molecular organization of the Ndc80 complex, an essential kinetochore component. Proceedings of the National Academy of Sciences of the United States of America *102*, 5363-5367.

Weir, J.R., Faesen, A.C., Klare, K., Petrovic, A., Basilico, F., Fischbock, J., Pentakota, S., Keller, J., Pesenti, M.E., Pan, D.Q.*, et al.* (2016). Insights from biochemical reconstitution into the architecture of human kinetochores. Nature *537*, 249-+.

Welburn, J.P.I., Vleugel, M., Liu, D., Yates, J.R., Lampson, M.A., Fukagawa, T., and Cheeseman, I.M. (2010). Aurora B Phosphorylates Spatially Distinct Targets to Differentially Regulate the Kinetochore-Microtubule Interface. Molecular Cell *38*, 383-392.

Westermann, S., Drubin, D.G., and Barnes, G. (2007). Structures and functions of yeast kinetochore complexes. Annual review of biochemistry *76*, 563-591.

Wittenberg, C., and Reed, S.I. (2005). Cell cycle-dependent transcription in yeast: promoters, transcription factors, and transcriptomes. Oncogene *24*, 2746-2755.

Wittmann, T., Hyman, A., and Desai, A. (2001). The spindle: a dynamic assembly of microtubules and motors. Nature cell biology *3*, E28-34.

Xu, D., Jaroszewski, L., Li, Z.W., and Godzik, A. (2014). FFAS-3D: improving fold recognition by including optimized structural features and template re-ranking. Bioinformatics *30*, 660-667.

Yamagishi, Y., Yang, C.H., Tanno, Y., and Watanabe, Y. (2012). MPS1/Mph1 phosphorylates the kinetochore protein KNL1/Spc7 to recruit SAC components. Nature cell biology *14*, 746-752.

Yanagida, M. (2014). The role of model organisms in the history of mitosis research. Cold Spring Harbor perspectives in biology *6*, a015768.

Yang, C.H., Tomkiel, J., Saitoh, H., Johnson, D.H., and Earnshaw, W.C. (1996). Identification of overlapping DNA-binding and centromere-targeting domains in the human kinetochore protein CENP-C. Molecular and Cellular Biology *16*, 3576-3586.

Zeitlin, S.G., Shelby, R.D., and Sullivan, K.F. (2001). CENP-A is phosphorylated by Aurora B kinase and plays an unexpected role in completion of cytokinesis. J Cell Biol *155*, 1147-1158.

6 Acknowledgements

I would like to acknowledge my supervisor Prof. Dr. Andrea Musacchio for the support. Thanks to the following people that contributed to the project or the helpful discussion: Dr. Ingrid Vetter, Dr. Dongqing Pan, Dr. Arsen Petrovic, Dr. Jenny Keller, Dr. Pim J. Huis in't Veld, Dr. Tanja Bange, Dr. Arthrur Porfetye, Dr. Marion Pesenti, Dr. Stefano Maffini, Dr Arnaud Rondelet, Dr. Charlotte Smith, Dr. Maria Thanasoula, Dr. Alex Faesen, Dr. Siva Jeganathan, Dr. Marta Mattiuzzo, Dr. Federica Basilic, Dr. Kerstin Klare, Dr. Claudia Breit, Dr. Giuseppe Ciossani, Dr. Valentina Piano.

And also the help from: Stefan Baumeister, Lin Yu-Chih, Kathrin Estel, Hendrik Hausmann, Ingrid Hoffmann, Carolin Koerner, Franziska Müller, Katharina Overlack, Satyakrishna Pentakota, Antje Peukert, Sabine Wohlgemuth, Annika Sarembe, Divya Singh, Patricia Stege, Isabelle Stender, Anika Take, Christiane Theiss, Dorothee Vogt, Beate Voß.

Special thanks to Christa Hornemann and Dr. Lucia Sironi for leading the IMPRS. Thanks all the people in MPI, Dortmund.

Max-Planck-Institut
für molekulare Physiologie

Curriculum vitae

The biography is not included in the online version for reasons of data protection.

Erklärung:

Hiermit erkläre ich, Yahui Liu gem. § 7 Abs. (2) d) + f) der Promotionsordnung der Fakultät für Biologie zur Erlangung des Dr. rer. nat., dass ich die vorliegende Dissertation selbständig verfasst und mich keiner anderen als der angegebenen Hilfsmittel bedient, bei der Abfassung der Dissertation nur die angegeben Hilfsmittel benutzt und alle wörtlich oder inhaltlich übernommenen Stellen als solche gekennzeichnet habe.

Essen, den 21.06.2017 Yahui

Unterschrift des/r Doktoranden/in

Erklärung:

Hiermit erkläre ich, Yahui Liu gem. § 7 Abs. (2) e) + g) der Promotionsordnung der Fakultät für Biologie zur Erlangung des Dr. rer. nat., dass ich keine anderen Promotionen bzw. Promotionsversuche in der Vergangenheit durchgeführt habe und dass diese Arbeit von keiner anderen Fakultät/Fachbereich abgelehnt worden ist.

Essen, den 21.06.2017 Yahui

Unterschrift des/r Doktoranden/in

Erklärung:

Hiermit erkläre ich, Prof. Dr. Andrea Musacchio gem. § 6 Abs. (2) g) der Promotionsordnung der Fakultät für Biologie zur Erlangung der Dr. rer. nat., dass ich das Arbeitsgebiet, dem das Thema „**Biochemical and Structural reconstitution of the outer Kinetochore of Drosophila melanogaster**" zuzuordnen ist, in Forschung und Lehre vertrete und den Antrag von Yahui Liu befürworte und die Betreuung auch im Falle eines Weggangs, wenn nicht wichtige Gründe dem entgegenstehen, weiterführen werde.

Essen, den 21.06.2017

Unterschrift eines Mitglieds der Universität Duisburg-Essen